VAXcluster System Management

Student Workbook

EY-9788E-SG.0004

Educational Services

digital™

VAXcluster System Management Student Workbook

EY-9788E-SG-0004

This document was prepared and published by Educational Services Development and Publishing, Digital Equipment Corporation.

CONTENTS

3 VAXcluster Software Concepts

4 Planning a VAXcluster System

5 Building a VAXcluster System

6 Managing VAXcluster Operations

EXAMPLES

FIGURES

TABLES

About This Course

INTRODUCTION

The VAXcluster System Management course is designed to teach experienced VMS system managers how to manage a group of VAX and MicroVAX systems running the VAXcluster software.

This student workbook is divided into seven modules, or units, each designed to cover a well-organized topic or group of topics. Most modules include figures, tables, and examples to enable you to better understand the material.

This course can be taught as either a lecture/lab or a seminar. If you are taking the lecture /lab version of this course, you will find a separate laboratory exercise book at the back of this workbook.

This module, *About This Course*, describes the contents of the course and suggests ways to use its materials most effectively. The following topics are discussed here:

- Course Description
- Prerequisites
- Course Goals
- Nongoals
- Resources
- Course Organization
- Course Map
- Course Conventions

COURSE DESCRIPTION

This course is for experienced VMS system managers. It teaches the purpose of a clustered environment, as well as how to plan and perform the functions needed to build, monitor, and maintain a VAXcluster system.

PREREQUISITES

To derive the greatest benefit from this course, you should know how to manage a nonclustered VMS Version 5 system. You can gain this knowledge by taking:

- Version 5.0, VMS System Management course

  OR

- Version 4.0, VMS System Management course and VMS Version 5.4 Technical Update

In addition, you should have at least six months of system management experience.

COURSE GOALS

This course is intended to give an understanding of:

- Characteristics of a VAXcluster system
- Purpose of the VAXcluster hardware components
- How the VAXcluster software functions
- Issues involved in choosing and planning a VAXcluster environment
- Issues involved in preparing to build a VAXcluster system
- Process and individual procedures necessary to build a VAXcluster system
- Procedures necessary to maintain and reconfigure a running VAXcluster system
- How to identify and solve VAXcluster problems

NONGOALS

This course does not address the following topics:

- Tuning the VAXcluster system
- Detailed VAXcluster specification and design
- VAXcluster application design
- Management of a nonclustered VMS system
- VAXcluster hardware and software internals

COURSE ORGANIZATION

This course is organized into a series of modules. Each module has its own learning objectives and covers a single topic or group of closely related topics. A module consists of:

- An **introduction**, which describes the purpose of the module, provides motivation for mastering its objectives, and outlines its contents.
- One or more **objectives**, which identify the skills taught in the module. Objectives are designed to focus your study efforts on a selected number of skills.
- The module **text**, which consists of:
  - Descriptive text organized in a list format
  - Illustrations, which clarify the relationships among various elements of a VMS system, or summarize steps of a particular process or command
  - Examples containing sample listings from actual interactive sessions on a VMS system
- A module **summary**, which reviews important concepts and skills taught in the module.

Written exercises are provided with this course. If you are taking the lecture/lab version, laboratory exercises are also provided in a separate booklet. These exercises help you review and practice the skills you learned during the lecture session.

RESOURCES

Students must have access to the following manuals to perform the recommended learning activities of this course. Students may be given their own copy of some of these manuals and the instructor may provide others for reference during the week.

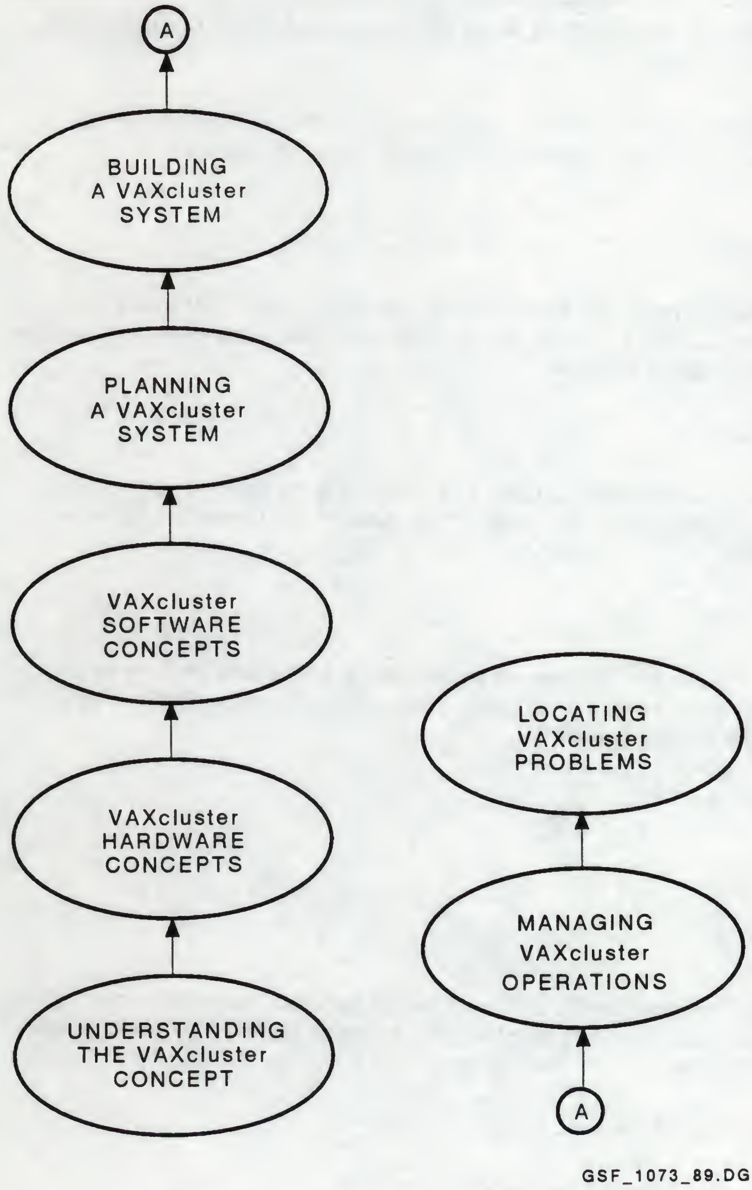
- *VMS VAXcluster Manual*
- *VAXcluster Software V5.4 Software Product Description (SPD)*
- *HSC User Guide*
- *VMS Show Cluster Utility Manual*
- *VAX Volume Shadowing Manual*
- *VMS SYSMAN Utility Manual*
- *VMS License Management Utility Manual*
- *Guide to Setting Up a VMS System*
- *Guide to DECnet–VAX Networking*
- *VMS Networking Manual*
- *VMS Network Control Program Manual*
- *VMS Version 5.4 Release Notes* (or most recent release notes)
- *VMS Monitor Utility Manual*
- *VMS System Services Reference Manual*
- *VMS Device Support Manual*

- *VAXcluster System and Application Design*
- *VMS DCL Dictionary*
- *VMS System Generation Utility Manual*
- *Introduction to VMS System Services*
- *Guide to VMS Performance Management*
- *VMS System Dump Analyzer Utility Manual*
- *Guide to Maintaining a VMS System*
- *VMS Access Control List Editor Manual*
- *VAXcluster Systems Handbook*
- *Introduction to VMS System Management*
- *Networks and Communications Buyer's Guide*
- *VAX Systems/DECsystems Systems and Options Catalog*
- *Guide to Using VMS Command Procedures*
- *Getting Started with VAXsimPLUS*
- *VAXsimPLUS User Guide*
- *Introduction to VAXcluster Application Design*
- *VMS Error Log Utility Manual*

At least one copy of the entire extended VMS documentation set should be available for reference.

COURSE MAP

This course map shows how each module of the course is related to the other modules and the course as a whole.



COURSE CONVENTIONS

Table 1 describes the conventions used in the listings and command tables of the Student Workbook.

Table 1 Course Conventions

Convention	Meaning
CTRL/X	Press and hold the key labeled CTRL while you press another key (X). Many control keys have special meanings.

UPPERCASE	In commands, uppercase characters indicate words you type exactly as they appear. For example, you would type the following commands as they appear:
	
	`$ DIRECTORY`
	`$ TYPE LOGIN.COM`

lowercase	Lowercase characters represent elements that you must replace according to the description in the text. For example, you must follow certain rules when you replace "file-spec" in the following example: `$ TYPE file-spec`

Ellipsis
(. . .)
	Horizontal ellipses indicate that you can enter additional parameters, values, or information. For example, you can enter any number of file specifications in the following example:
	
	`$ TYPE file-spec, . . .`
	
	Vertical series of periods or ellipses mean that not all of the data that the system would display in response to the particular command is shown, or that not all the data a user would enter is shown.
	
	`$ TYPE MYFILE.DAT`
	`.`
	`.`
	`.`
	`$`

Square brackets ([ ])	Square brackets indicate that the enclosed item is optional. (Square brackets are not optional, however, in the syntax of some file specifications.) For example, the logical name is optional in the following command: `$ MOUNT/FOREIGN $TAPE1 [logical-name]`
Braces ({})	Indicate that you must select from the included items.
Quotation Marks and Apostrophes	The term quotation marks refers to double quotation marks ("). The term apostrophe refers to a single quotation mark (').

Understanding the VAXcluster Concept

INTRODUCTION

This module introduces the VAXcluster concept and what it means to connect several VAX processors into a single cluster. It compares a VAXcluster system with a single, nonclustered VAX system and with a network of systems. It also describes the different types of clusters available.

OBJECTIVES

After completing this module, students should understand:

- VAXcluster concepts
- VAXcluster components
- DECnet VAX communication in a VAXcluster system
- The major differences between single VMS systems and VAXcluster systems

RESOURCES

- *VMS VAXcluster Manual*
- *Introduction to VMS System Management*
- *VAX Systems/DECsystems and Options Catalog*
- *Networks and Communications Buyer's Guide*

THE VAXcluster CONCEPT

A **VAXcluster system**

- Consists of cooperating nodes
- Uses VMS software
- Is connected by a high-speed communication medium in order to share data and resources

Members of a Cluster

- Share resources
  - Processing resources
  - Queues
  - Disk storage
- Use a single VMS security and management domain
- Boot and fail independently
- Communicate with all other VAXcluster processors to coordinate cluster activities
- Potentially perform I/O to any disk storage subsystem in the cluster
- Execute a private copy of VMS in memory

Common Environment VAXcluster Systems

- VMS nodes share system files, generally using the same system disk.
- Use the same resources
- Have identical user accounts on members
- Access the same images
- Define the same logical names

Multiple Environment VAXcluster Systems

- Share one set of resources
- Designate individual nodes to perform specific functions

VAXcluster COMPONENTS

A VAXcluster system is made up of all or some of the following components:

- VAX processors (VMS members, active nodes)
- Storage Controllers and Storage Devices
  - HSC (Hierarchical Storage Controller) units in clusters using the CI bus
  - Integrated Storage Elements (ISEs) in clusters using the DSSI bus
  - Other storage devices, connected to VAX processors
- Interconnects

VAX Processors

VAX or MicroVAX processors running the VMS operating system may participate in a VAXcluster system. Any VAX processor in a cluster is called an **active node** or a **VMS member**.

- There is no master node in a VAXcluster system.
- Each VMS system is a full member of the cluster.
- Parameters can be adjusted to encourage large, fast nodes to do most of the upkeep of cluster databases.
- Parameters can be adjusted to minimize the disruption caused by nodes that frequently must leave the cluster.

VMS Members

Members can share the following:

- System disks
- Data disks
- Queues
- Applications
- User authorization files
- LAT based printers
- Tape drives

Local to each member can be the following:

- Memory
- Tape drives
- Printers
- Disks

Specific to each member are the following:

- Page/swap files
- System parameter files

Special Members

Certain members in the cluster can be designated to carry out special functions or responsibilities.

Boot Servers = Boot member = Boot node

- Down-line load satellite boot files
- Serve the system disk to satellites
- Contain:
  - Cluster common files for startup
  - Directory roots that boot satellites

Satellites

- Are MicroVAX or VAXstation members
- Do not have local system disks
- Boot remotely over the Ethernet from the boot server node

Disk Servers

- Allow other VAXcluster nodes access to disks to which those nodes do not have a direct connection
- Can serve:
  - Locally attached disks
  - The HSC disks if connected to an HSC
  - The DSSI disks if connected to a DSSI

Ethernet Members

- Use local system disk
- May serve local disks to other members, thus acting as a disk server

HSC (Hierarchical Storage Controller) and DSSI Integrated Storage Element (ISE) Nodes

- Do not participate in cluster negotiations
- Unaware of locking
- Read and write to disks as requested by cluster members
- HSC nodes can provide cluster-accessible tape drives

Interconnects

The type of VAXcluster configuration is determined by the interconnect device used for System Communication Services (SCS) communication.

CI (Computer Interconnect) Bus blauw

max 45m

- High-speed (70 megabits/sec), serial dual-path bus max 16 nodes
- Connects VAX processors and HSC units, in conjunction with the Star Coupler and the VAX and HSC CI port controllers.

Ethernet geel of grijs

- Lower speed (10 megabits/sec), serial single-path bus
- Multiple protocols can be transmitted DECnet transmissions and interprocessor SCS in some configurations cluster verkeer

DSSI (Digital Storage Systems Interconnect) Bus grijs + groen/geel aardkabel

- Short (~~6~~ 25 meters), fast (4 megabytes/sec), parallel bus max 8 nodes (disken zijn ook nodes)

Mixed Interconnect

- Any combination of the three types of interconnect

See the most recent VAXcluster Software Product Description (SPD) for specific configuration rules.

CI (Computer Interconnect) Based VAXcluster System

The CI VAXcluster system is a set of VMS members that communicate using only a CI (Computer Interconnect) bus for cluster communications.

It features:

- The Star Coupler as a common connection point for all cluster nodes
- CI redundancy that enhances availability
- HSC (Hierarchical Storage Controller) units that free members from some I/O processing and optimize disk access cluster-wide

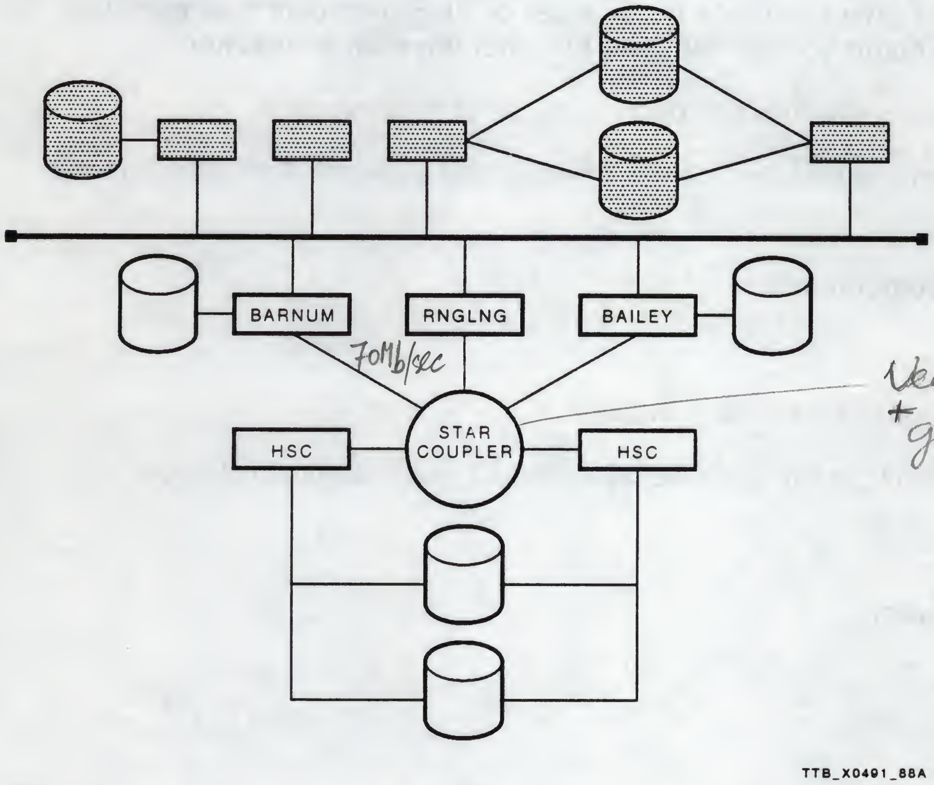
CI VAXcluster System Advantages:

- Largest storage capacities and transaction rates available in clusters
- Fully redundant hardware for highest availability
- CI service for up to 32 connections 16 active 16 passive
  - VMS systems (members) or HSC storage subsystems
  - Minimum of one VMS system
- HSC or Host-based volume shadowing available
- Centralized system management
- Can be accessed through DECnet and act as a network server of cluster disks and other resources

CI VAXcluster System Disadvantages:

- Limited distance between CI nodes
- Cost
- Geen µVAX mogelijk

Figure 1–1 Typical CI VAXcluster System with HSC units



verbindings apparaat
+ galvanische scheiding

The unshaded area of this figure shows a typical CI VAXcluster system with HSC units:

- BARNUM = VAX 8820
- RNGLNG = VAX–11/785
- BAILEY = VAX 6210
- HSC = HSC40, HSC50, or HSC70
- Disks are RA60, RA70, RA82, RA90

Ethernet-Based Local Area VAXcluster System

A local area VAXcluster system is made up of a set of VMS members that establish cluster connections and communicate using an Ethernet physical connection:

- MicroVAX and VAXstation satellites that boot by means of the Ethernet
- Boot servers that down-line load the operating system and allow satellites access to their system disk

Local Area VAXcluster Components:

- Boot server
  - Management center and resource provider
  - System disk contains the system directory roots and cluster common files for:

    Startup

    User authorization

    Queue setup
  - Makes available:

    Data disks

    Printers

    Distributed batch processing
- DECnet
  - DECnet Maintenance Operation Protocol (MOP) responds to satellite requests for downloading of VMS system
  - Sends VMS image to node
- Satellites are generally consumers (not providers) of cluster resources
  - Boot remotely from a system disk on the bootserver
  - Can use local disks for paging and swapping to improve performance

nu NI

Local Area VAXcluster System Advantages:

- Relatively low cost for VAXcluster benefits.
- Cluster service for up to 96 members.
- Hardware is also used for networking.
- Ability to have wide geographic distribution.
- Several separate clusters can coexist over the same Ethernet.
- Disk space optimization versus each satellite storing identical system files.
- Centralized system management, instead of individually managed systems.
- Large VAX boot servers can offer high-performance batch queues.
- Using dual-hosted system disks, multiple boot servers, and a quorum disk, the cluster can be configured to be highly available.
- Host-based disk shadowing is available.

nu NI

Local Area VAXcluster System Disadvantages:

- I/O throughput is limited by Ethernet adapters and Ethernet traffic. 10 Mb/sec
- Network Interconnect SCS(NI-SCS) is more CPU intensive than SCS.
- There is a possibility of system disk bottleneck.
- The Ethernet cable is a single point of failure.

DSSI (Digital Storage Systems Interconnect) Bus

The DSSI bus allows certain processors and Integrated Storage Elements to communicate. More than one processor can reside on the bus, thereby providing more than one path to the ISE if the processors are also on an Ethernet.

This dual-host configuration also allows VMS MicroVAX 3300 and MicroVAX 3400 members to communicate using the DSSI bus for disk access and cluster (SCS) communications.

It features:

- The DSSI bus
- The Integrated Storage Element (ISE) which acts as both controller and storage: voor RF schijven
  - Provides essential features of an HSC subsystem
  - Provides mass storage
- The KFQSA and Embedded DSSI Adapter (EDA) for DSSI bus to MicroVAX connections
  - EDA embedded onto MicroVAX 3300 and MicroVAX 3400 CPU module
  - KFQSA adapter connects DSSI bus to Q-bus voor MVAX II noodzakelijk

SHAC ook 4300

Dual Host Configuration Advantages:

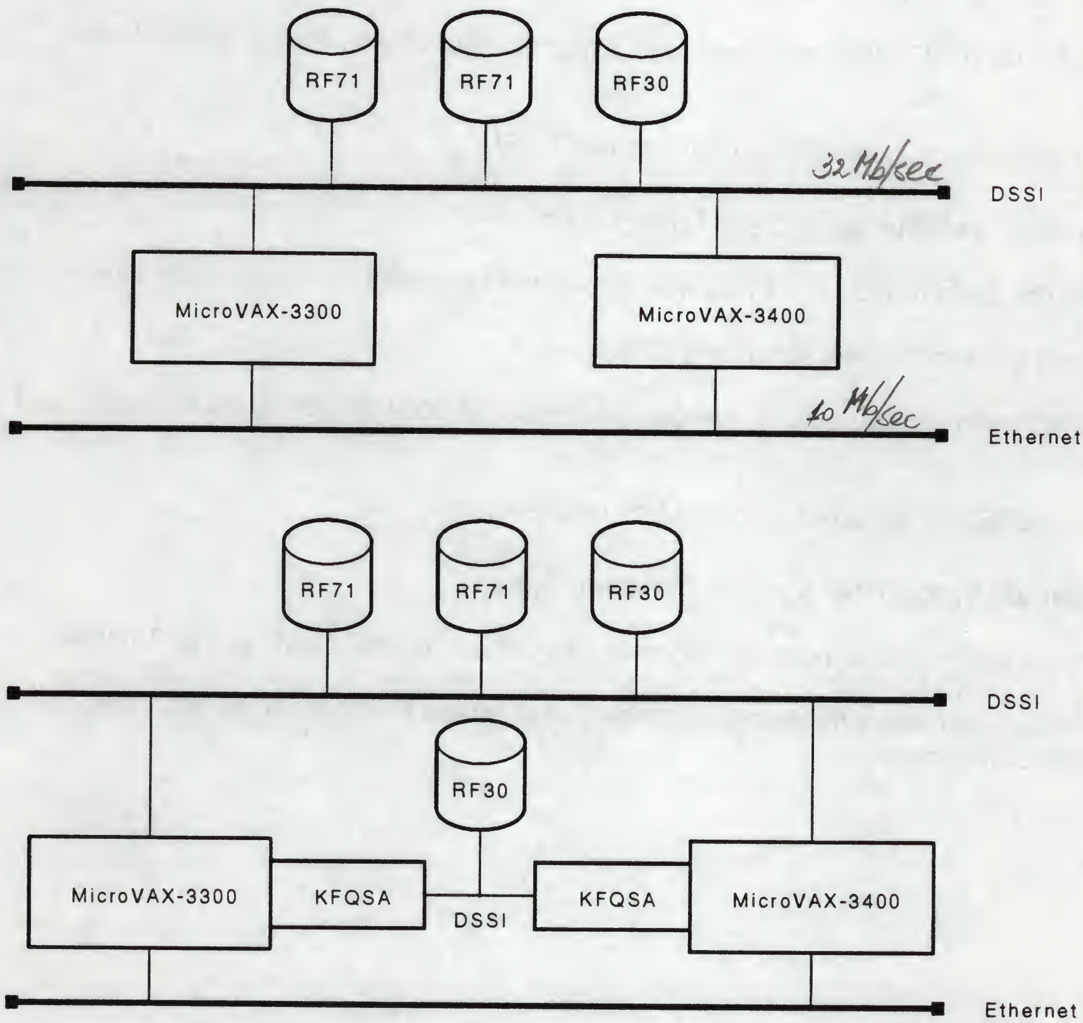
- Combines storage and controller functionality
- Eliminates contention for controller resources resulting in better I/O performance
- Centralized system management
- Lower cost

Dual Host Configuration Disadvantages:

- Limited distance between DSSI nodes (25m)
- Limited number of host systems supported max 3 bootservers

Satelieten 96

Figure 1–2 Typical DSSI VAXcluster System Using EDA and KFQSA Adapters



GSF_1071_89.DG

Mixed-Interconnect VAXcluster System

The Mixed-Interconnect VAXcluster system is made up of a set of VMS members that communicate over any combination of CI, DSSI, and Ethernet interconnects.

It features:

- CI or DSSI members that can act as boot servers for satellites
- CI service for up to 32 VMS members and Ethernet (NI) service for up to 96 cluster members
- DSSI service for up to 2 MicroVAX members per DSSI

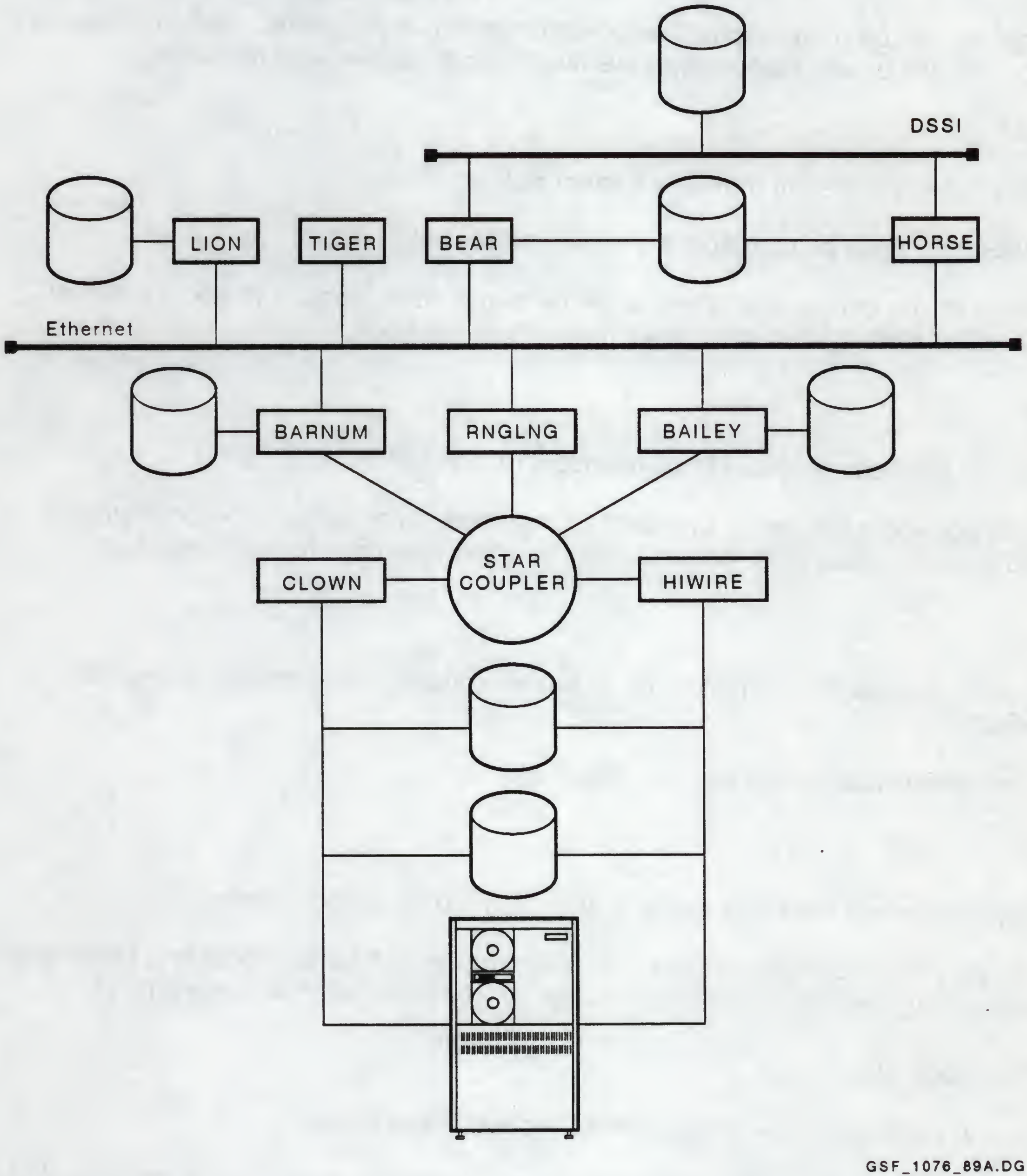
Mixed-Interconnect VAXcluster System Advantages:

- Combines high performance of CI bus with distributed convenience of local area cluster
- Allows higher availability than local area cluster
  - Satellite system disks on either dual-ported, volume shadowed HSC disks or dual-hosted DSSI disks
  - All CI or DSSI nodes can be enabled as boot servers

Mixed-Interconnect VAXcluster System Disadvantages:

- Care must be taken that CI boot servers and disk servers do not overload the Ethernet.
- Care must be taken that satellite members do not overload the boot server and disk server nodes with I/O requests.

Figure 1–3 Typical Mixed-Interconnect VAXcluster System



This figure shows a VAXcluster system that combines features of the types of interconnects discussed in this module.

Resource Sharing in a VAXcluster System

System resources can be made accessible to each member of the cluster. Each member can use the resources just as any standalone VMS member can access local resources.

System disks

- Each member has a system root on a system disk.
- Many VMS nodes can all boot from the same system disk.
- All systems in the cluster may share a single system disk (subject to some practical restrictions), or each system could have its own system disk.

Data disks

- All disks can be made available to all members of the cluster.
- Disks can physically be ported to individual members, dual-ported between members, attached to HSC nodes, or in the case of ISEs, dual-hosted by MicroVAX members.

Queues

- Cluster-wide queues are controlled by a cluster common job controller queue file, JBCSYSQUE.DAT.
- Users can submit jobs to any queue in the cluster.

Applications

- Most applications will work in a cluster just as they did on a single system.
- Applications can be designed to take advantage of the distributed processing capabilities of clusters, using the distributed lock manager for communication and coordination.

User authorization files

- All nodes in the cluster can use the same user authorization file.
- All user passwords, directories, limits, quotas, and privileges could then be the same on all systems.

Resources Local to Each Node

Memory

- Each cluster member maintains its own memory.
- This prevents multiple systems from failing when one node develops problems with consistency of data in its memory.
- Any node can leave the cluster at any time without necessarily causing the cluster as a whole to fail.

Tapes

- Local to a single system at a time.

Devices

- Other devices are considered local to a specific member.
- However, any device that accepts input through queues can be used from any node in the cluster.
- The cluster Job Controller will allow any node to submit a job to any queue in the cluster.

Resources Specific to Each System

Page and Swap files

- These can reside on any disk. Thus, while specific to a particular system, they need not be on that system's system disk. They may reside on a local disk for better performance.

System parameters

- The various system parameter files reside in that system's root directory on the system disk.

DECnet VAX COMMUNICATIONS IN A VAXcluster SYSTEM

Any cluster configuration requires DECnet VAX communications for all processor nodes.

- Ensures access to all nodes from a single terminal
- Local area VAXcluster and Mixed Interconnect (MI) systems
- Required for system management (SYSMAN)
- Required for interprocessor communications
- DECnet and SCS software coexist on nodes on the Ethernet
- DECnet and SCS use same data link and physical protocols

COMPARISONS BETWEEN VAXcluster AND SINGLE VMS USER ENVIRONMENTS

An important aspect of the VAXcluster concept is that the user environment is the same as on a single VMS system. If you configure a cluster so that it provides a common user environment, then users need no knowledge of the cluster configuration — not even which node they are logged in to — to do their work.

Table 1–1 A Comparison of VMS Features Between Standalone Systems and VAXcluster Common User Environments

Feature	Standalone VMS System	VAXcluster System
Disk Sharing	Processes can share VAX RMS files on the system. Read and write access can be shared to the record level.	Processes can share VAX RMS files anywhere on the cluster. Read and write access can be shared to the record level.
Job Control	System users can submit jobs to any batch or print queue on the system.	Cluster users can submit jobs to any queue in the cluster, regardless of the system on which the queue will actually execute. Generic queues can balance the load among the available processors.
Lock Management	Locks are managed among the processes on the system.	Locks are managed among the processes in the cluster.
User Environment	A user can customize the DCL environment so that it is available whenever he or she logs in to the system.	A user can customize the environment so that it is available whenever he or she logs in to any member of the cluster.
System Management	The VMS system manager has access to tools that control the user population, disk configuration, and system parameters.	The VAXcluster system manager has access to tools that centrally control the user population, disk configuration, and system parameters of all members in the cluster.

In comparison, networked systems are communicating single systems. They need to know about each other in terms of routing issues and resources or services that may be provided across the network. But they remain individual management domains which must be coordinated to provide services throughout the network and to maintain secure systems.

Multiprocessors share the same memory and are capable of sharing images. Multiprocessor applications can take advantage of interprocess mechanisms (interprocess communication, common event flags, etc.) to coordinate activities. Multiprocessor systems may participate in networks and clusters.

Table 1–2 Comparing Multiprocessor Systems

System Characteristic	Network	VAXcluster System	VAX 6000 Multiprocessor
CPU location	Local or wide area	Same computer room (CI) or local area	Single or adjacent
Security/management domain	Multiple	Single	Single
File system	Separate	Integrated	Integrated
Operating system	Separate (some may not be VMS systems)	Separate (all VMS systems)	Shared
Growth potential	Very great	Somewhat limited	Limited

SUMMARY

- A VAXcluster system consists of a number of cooperating nodes, which can be:
  - VAX processors
  - MicroVAX processors
  - Storage subsystems
- VAXcluster systems use VMS software and are connected by a high-speed communication medium in order to share data and resources.
- The components of a VAXcluster system are:
  - VAX processors
  - Storage controllers
  - Interconnects
  - Mass storage devices
- Some major differences between standalone VMS systems and VAXcluster systems are:
  - Disks and files are accessible from any node in a cluster.
  - Jobs can be submitted from any node to any queue in a cluster.
  - Resource access is managed by a cluster-wide lock manager.
  - Users can customize the environment to be available on any node in the cluster.
  - System and user management can be handled at the cluster level.

VAXcluster Hardware Concepts

INTRODUCTION

This module introduces the basic hardware components of a VAXcluster system.

As a system manager, you may be responsible for choosing hardware components and deciding how to configure them into the cluster. You may design your own cluster or purchase a prepackaged VAXcluster system. Once you have formed your cluster, you may want to expand it as your computing needs increase.

OBJECTIVES

After completing this module, students should be able to describe:

- VAX processors
- HSC (Hierarchical Storage Controller) units
- ISE (Integrated Storage Element) units
- The three types of interconnects and their components:
  - Ethernet
  - CI (Computer Interconnect) bus
  - DSSI (Digital Storage Systems Interconnect) bus
- Mass storage components
- VAXcluster console system
- Prepackaged VAXcluster systems

RESOURCES

- *VAX Systems/DECsystems Systems and Options Catalog*
- *Networks and Communications Buyer's Guide*
- *VMS VAXcluster Manual*
- *Introduction to VMS System Management*
- *VAXcluster Systems Handbook*

VAXcluster HARDWARE

You can form a VAXcluster system from just a few components. Your VAXcluster system may start as small as one or two VAX or MicroVAX systems and may or may not have an HSC unit. The cluster can include new systems, as well as most of your existing systems. Its interconnect can be a CI or DSSI bus, Ethernet, or any combination of the three.

Once you have formed a cluster, you can expand it as your computing needs increase:

- As your need for computing power increases, you can increase the number of VAX and MicroVAX systems in your cluster.
- As your mass storage needs increase, you can add more storage controllers and devices.
- As the complexity of the cluster increases, you can incorporate two or more types of interconnect.

VAX PROCESSORS

The following processor families are supported in VAXcluster configurations:

- VAX family
- VAX–11 family
- MicroVAX family
- VAXstation family

VAX Members

- All VAX members may be boot members in a local area VAXcluster system. NI
- There can be no more than one VAX system more powerful than a VAX 8350 in a local area VAXcluster (maximum one of VAX 9000, VAX 6000, VAX 8500, VAX 8600, VAX 8700, and VAX 8800 series).
- In a mixed-interconnect cluster, all members must have Ethernet connections.
- These configuration rules are subject to change.

VAX 9000 series

The VAX 9000 series (and VAX 6000-400) uses the XMI bus (80 megabytes/sec)

- XMI disk controller: KDM70
- XMI Ethernet port: DEMNA
- XMI CI port: CIXCD

VAXBI series

These systems use the high-performance VAXBI bus for most I/O devices

- VAX 8700, VAX 8800 series
- VAX 6000 series
- VAX 8530, VAX 8550 systems
- VAX 8250, VAX 8350 systems
- VAXBI disk controller: KDB50
- VAXBI Ethernet port: DEBNA
- VAXBI CI port: CIBCA

UNIBUS and MASSBUS VAX series

These systems use the MASSBUS and UNIBUS buses for most I/O devices

- VAX 8600, VAX 8650 systems
- MASSBUS device controllers
- UNIBUS Ethernet port: DELUA
- SBI CI port: CI780

VAX–11 Members

- VAX–11/780 system
- VAX–11/785 system
- VAX–11/750 system

In local area and mixed-interconnect configurations:

- VAX–11 members must boot from a local system disk, if not booting across a CI bus.
- VAX–11 members can be boot servers.

VAX–11 systems use the UNIBUS bus for most I/O devices

- UNIBUS CI port for VAX–11/780 and VAX–11/785: CI780
- UNIBUS CI port for VAX–11/750: CI750
- UNIBUS Ethernet port: DELUA or DEUNA

MicroVAX Members

These MicroVAX members are supported in cluster configurations:

- MicroVAX II: Q-bus
- MicroVAX 3500 and MicroVAX 3600: Q-bus
- MicroVAX 3300 and MicroVAX 3400: Q-bus
- MicroVAX 3800 and MicroVAX 3900: Q-bus
- MicroVAX 2000

MicroVAX members can be:

- Boot servers in a local area or mixed-interconnect VAXcluster system
- Satellites in a local area or mixed-interconnect VAXcluster system
- Dedicated print servers or other specialized processors

Q-bus options

- KDA50 disk controller (for RA disks)
- DELQA, DESQA, DEQNA: Q-bus Ethernet ports
- RQDX3 disk controller (for RD disks)
- TQK50 tape controller (for TK tapes)
- RRD50 compact disk reader
- KFQSA Q-bus to DSSI interface

MicroVAX 2000 option

- DESVA Ethernet port

VAXstation Members

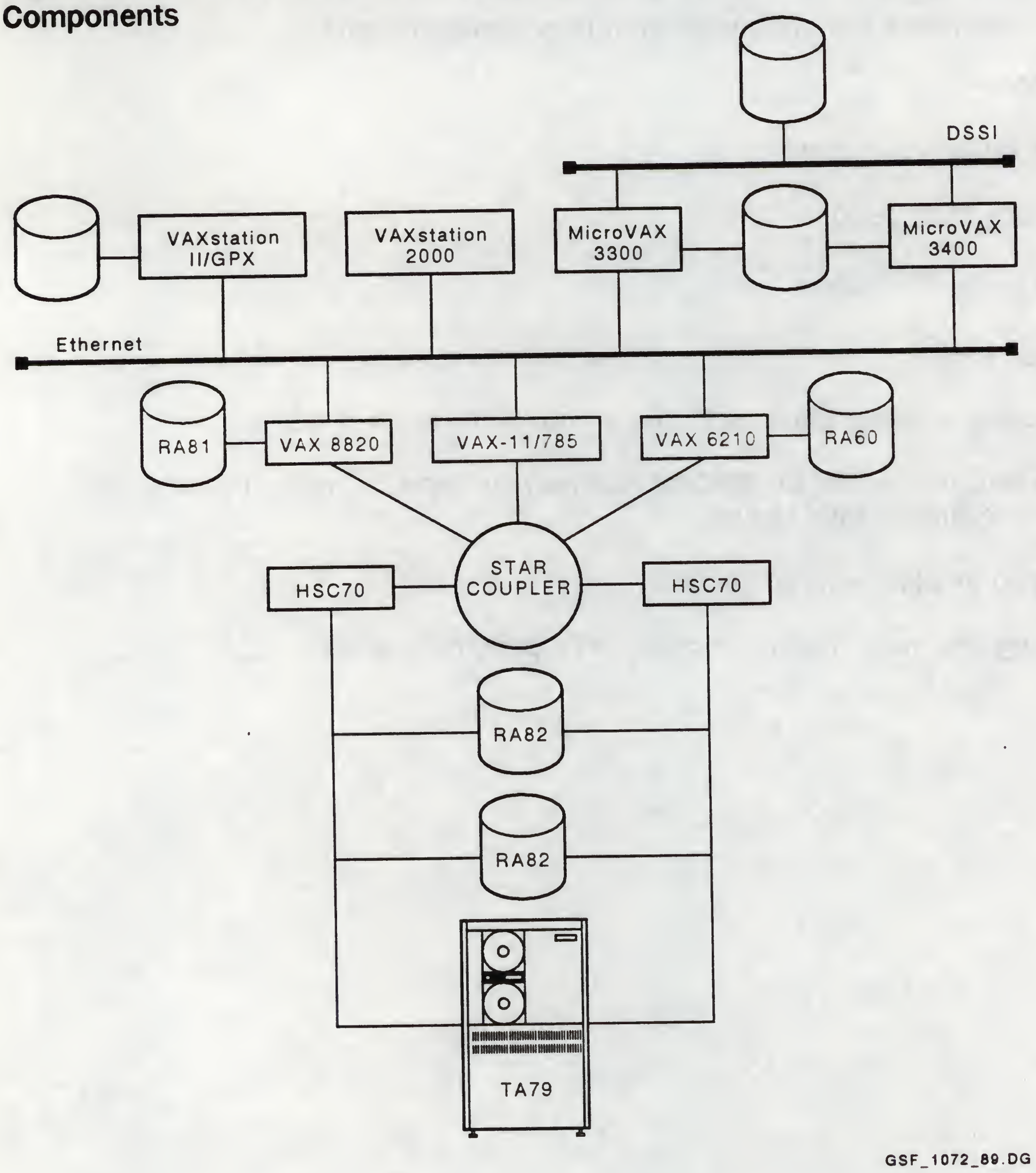
These VAXstation members are supported in cluster configurations:

- VAXstation 2000
- VAXstation VAXstation II, II/GPX
- VAXstation 3100, 3200, 3500
- AI VAXstation 2000, 3200, 3500

VAXstation members can:

- Be a boot server or a satellite in a local area or mixed-interconnect cluster
- Communicate with the cluster by DECnet and network products without having the VAXstation as a member of the cluster
- Allow bit-mapped graphics environments
- Allow processing and many system functions to be performed locally

Figure 2–1 A VAXcluster Configuration Using the Various VAXcluster Hardware Components



Processor Configuration Issues

There are many issues that affect the configuration of processors in a VAXcluster system. When you plan your CPU configuration, consider whether your interconnect is to be CI, Ethernet, DSSI or mixed-interconnect. Also consider the distribution of the target user population over the cluster. Among the considerations are:

- Available hardware. If possible, you want to make use of the Ethernet or CI hardware you have already.
- CPU types and locations. MicroVAX systems can be connected only to Ethernet. CI cables can be no longer than 45 meters, confining a CI only cluster to a single computer room or adjacent rooms. DSSI cables can be no longer than 6 meters, and can only presently be directly connected to Q-bus systems .
- Present and future expansion needs. Currently, there can be only 32 CI nodes (using Star Coupler Expander hardware), and there is a limit of two VMS nodes on a DSSI bus. The maximum number of VMS nodes in a local area or mixed-interconnect VAXcluster system is 96. (See the VAXcluster SPD for current restrictions.)
- Availability of the cluster to users. The CI bus provides hardware redundancy, while Ethernet does not.

  The CI cluster is the most highly available cluster configuration.

  A dual-hosted DSSI configuration is also highly available.

  A mixed-interconnect cluster can be configured as highly available as the CI bus, but the Ethernet is still a single point of failure.

  A local area cluster can be configured with dual boot servers, dual-ported and/or dual-hosted disks, and a quorum disk to enhance availability.

  You must decide how much redundancy is necessary for your application. Although a VAXcluster configuration is highly available, it does not guarantee nonstop processing or absolute fault tolerance.
- The expected load on the cluster I/O subsystems and interconnect. The CI bus can sustain higher I/O rates than Ethernet, and the Ethernet capacity may need to be shared with other systems that are not part of the cluster.
- Cost. Consider not only the cost of the interconnect, but the expense of installing it.

INTERCONNECTS

The physical interconnect allows the VAXcluster software to communicate across all nodes in the cluster. Three physical links are supported:

- Ethernet
- CI (Computer interconnect) bus
- DSSI (Digital Storage Systems Interconnect) bus

Ethernet

Ethernet connects a number of systems into a local area network in which a VAXcluster system and other types of communication can take place.

- A VAXcluster system that uses Ethernet alone for cluster communication is a Local Area VAXcluster system.

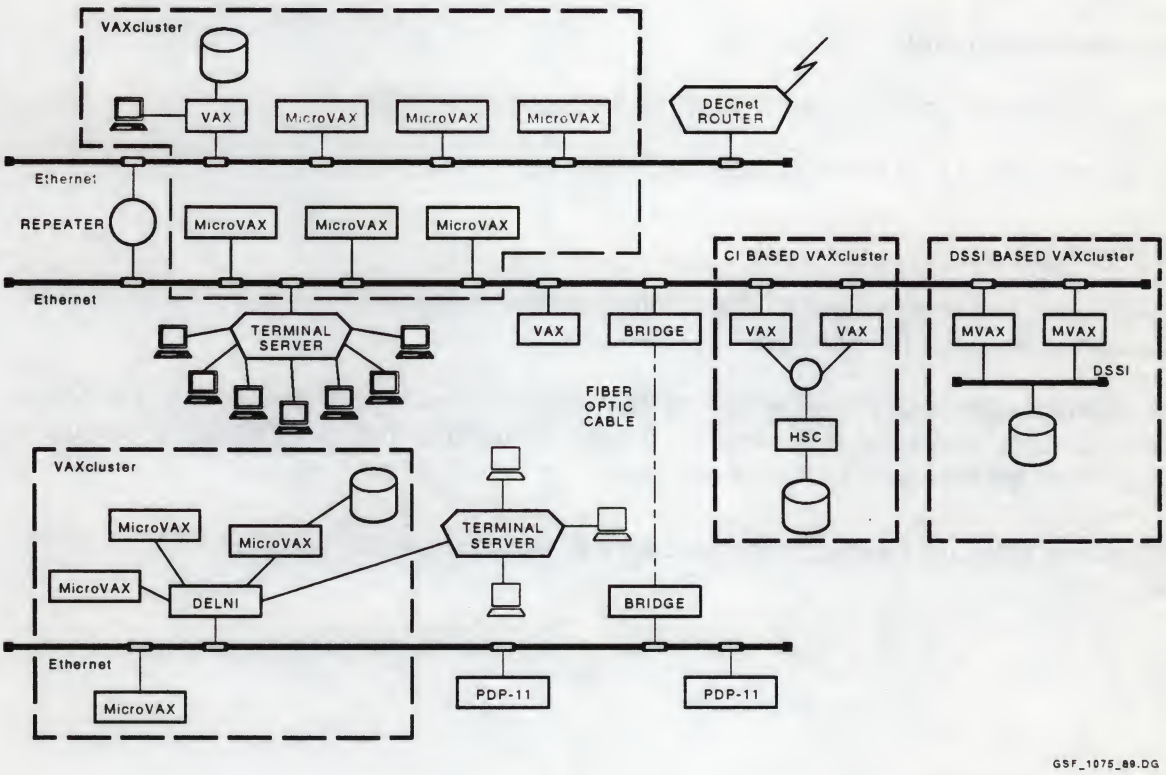
Ethernet features:

- Same hardware as used in local area networks
- 10 Mbits/sec bandwidth
- Configuration distances limited to less than 1 second propagation delay
- Repeaters and bridges that can be used to configure large and efficient network systems can be used for VAXcluster communications, as long as they meet the above two conditions

This section on Ethernet will discuss:

- Ethernet Hardware
- Ethernet Ports
- Ethernet Protocols
- Local Area Transport

Figure 2–2 Standard Ethernet Components for VAX and MicroVAX Systems, Terminal Servers, Bridges, and Repeaters



This figure shows an Ethernet network with some of the possible components. Each of the three horizontal Ethernet cables is called a **segment**. The entire configuration functions as a single logical Ethernet supporting several VAXcluster systems in addition to other systems and services.

Ethernet Hardware

There are three types of Ethernet cable that can be used for cluster communication:

Standard (thick wire) baseband cable

- Coaxial cable intended for communication within a single building or group of buildings.
- New hardware can be added, without interrupting network traffic, by tapping into the cable.

ThinWire baseband cable

- Intended for communication within a set of rooms or office areas
- Can be connected to a building's standard Ethernet cable

Broadband cable

- Allows video and other mixed-media communication protocols to share the same physical wire (cable television for example)
- Can connect baseband Ethernet segments, which can each have members in the same cluster, if traffic is modest and bridges are carefully used so the Ethernet is not overloaded and a 10 megabit/second path is guaranteed

The VAXcluster software functions independently of the physical configuration and the type of Ethernet cable.

Ethernet Ports

The Ethernet ports are the VAXcluster members' interface to the Ethernet.

- Support multiple protocols — each protocol is handled by a separate software module, and they are relatively independent of each other.
- Decode messages that are sent to the port hardware address.
  - Address is set by chips in the port board.
- Perform DMA transfers.

Some ports available:

- DESVA for MicroVAX and VAXstation 2000 systems
- DESQA, DELQA, DEUNA for Q-bus systems
- DELUA for UNIBUS systems
- DEBNA for VAXBI systems
- DEMNA for XMI systems

Ethernet Protocols

Each Ethernet protocol is layered directly onto the Ethernet hardware protocol.

A single Ethernet can support a large number of independent protocols simultaneously. Those protocols include:

- DECnet protocols, for a variety of operations including remote file access and down-line loading of software
- LAT (Local Area Transport) protocol, for communication between terminal services and host systems
- System Communication Services (SCS) protocol, for communication between VAXcluster members

LAT (Local Area Transport)

Terminal and Device Servers:

- Allow terminals to connect to any member offering service on the Ethernet
- Allow terminals to connect to a cluster without specifying a particular member, if the members of the cluster all offer the same service
  - LAT software load is balanced between available cluster (or network) members offering the same service.
- Allow LAT terminal lines to be used to connect cluster-wide or network-wide printers or other terminal-line devices, thus providing access to these devices from any node in the local area network

CI (Computer Interconnect) Hardware

The CI (Computer Interconnect), a communication link between VAXcluster nodes (both active and passive), has the following parts:

- CI bus hardware
- VAX CI ports
- Star Coupler
- HSC (Hierarchical Storage Controller) units

CI Bus Hardware

The CI bus is used to link processors and storage controllers to the Star Coupler.

- A high-speed, multiaccess bus with 70 Mbits/sec bandwidth on each path
- Consists of four (4) coaxial cables (two transmit, two receive)
  - — Maximum cable length — 45 meters (148 feet)
  - — Two paths for data transmission
  - — Data path selected randomly for each transmission by the CI port
- Dedicated to VAXcluster communication
- Can also be used for DECnet communication

The CI port has collision detection hardware.

Star Coupler

The Star Coupler is a central, passive connection point for CI cables in a VAXcluster system.

- All CI nodes (both VAX and HSC) in a VAXcluster system must be able to communicate with each other through some Star Coupler (not necessarily a single Star Coupler).
- The Star Coupler accommodates CI cables in a radial arrangement.

VAX CI Ports

The VAX CI Ports are microcoded, intelligent controllers that are connected to both paths of the CI bus and are used for connecting processors to the CI bus.

- Available for all VAX 9000 series, VAX 8800 series, VAX 6000 series, VAX–11/785, VAX–11/750 VAX–11/780, and systems
- Uses any available randomly chosen bus path
- When either path becomes unavailable:
  - — Traffic automatically shifts to the other path.
  - — Traffic is restored to the unavailable path when it becomes available again.
- Performs Direct Memory Access (DMA) for efficient mass storage access
- For systems with multiple CI ports, the load is automatically balanced across the ports.
- The CI port and the CI bus are optimized to perform large block transfers very efficiently with a particular, fixed-length packet size.

HSC (Hierarchical Storage Controller) Unit

The HSC is an intelligent disk and tape controller that performs functions required to manage and control the storage and transfer of information.

Components of an HSC storage subsystem are:

- Hierarchical storage controller (HSC40, HSC50, or HSC70)
- Digital Storage Architecture (DSA) disk and tape drives (RA disks, TA tapes)
- A console terminal

An HSC unit in a VAXcluster system:

- Is a passive node, but not a cluster member
- Contains a port for a single, dual-path CI bus
- Provides processes on any VAXcluster processor with access to devices connected to the HSC unit
- Is an MSCP server
- Optimizes request service to disk devices
- Contains resident diagnostics and utility programs
- Handles multiple, simultaneous operations on several drives
- Has up to three channels (HSC40), six channels (HSC50), or eight channels (HSC70) for connecting disks and tapes
  - Physical limit of four disks per channel
  - Physical limit of 16 tapes per channel
- Allows dual-porting of disks between two HSCs to improve disk availability
- Allows use of VAX Volume Shadowing (Phase I) software with the purchase of the license (controller-based volume shadowing)

HSC60 7 channels
HSC90 12 channels

DSSI (Digital Storage Systems Interconnect) Bus

The DSSI (Digital Storage Systems Interconnect) bus supports communication between Q-bus systems and Integrated Storage Elements.

The DSSI bus has the following parts:

- DSSI bus hardware and protocol
- Integrated Storage Element (ISE)
- Embedded DSSI Adapter (EDA)
- KFQ Storage Adapter (KFQSA)

The DSSI bus supports cluster communication between MicroVAX 3300 and MicroVAX 3400 host nodes using EDA640 adapters.

DSSI Hardware Protocol

- Provides a single 8-bit parallel multidrop data path with both byte parity and packet EDC
- Peak bandwidth — 32 Mbits/sec (>28 Mbits/sec usable)
- Maximum cable length — ~~6~~ 25 meters
- Supports up to 8 nodes, two of which may be host systems
- Low bus and node interface overhead

Integrated Storage Elements (ISEs)

The DSSI storage device, the RF-series Integrated Storage Element (ISE), combines the essential features of an HSC unit with mass storage capability.

- Each disk has its own controller, eliminating contention between multiple Head Disk Assemblies (HDAs) for a single controller.
- Each HDA, therefore, has its own dedicated MSCP server.

Features shared with HSC units:

- Serves itself to each host with simultaneous communication
- Communicates with the disk driver
- Handles physical placement of data
- Allows a high degree of data integrity
- Performs bad block replacement

ISEs do not offer the following HSC features:

- Controller-based volume shadowing
- Controller-based backup
- Dual-porting of disk ISEs (because disk and controller are one unit).

Dual-hosting of ISEs on the DSSI bus can be provided by Q-bus systems using any combination of EDA or KFQSA adapters, thus providing two paths to the ISE. This means that:

- Two systems on the same DSSI bus with an RF disk may directly access that disk.
- The dual-host configuration provides ISE failover for RFxx served to the rest of a local area or mixed-interconnect VAXcluster system.
- If one of the two systems on a DSSI bus serving an RF disk fails, the other will provide a path. Failover will be automatic.

Embedded DSSI Adapter (EDA)

The EDA is used by MicroVAX 3300 and MicroVAX 3400 systems to connect these members to the DSSI bus.

- Built into MicroVAX 3300 and MicroVAX 3400 systems
- Supports host-to-host cluster (SCS) communication across the DSSI bus.

KFQ Storage Adapter (KFQSA)

The KFQSA connects the DSSI bus to a MicroVAX system. It attaches to the Q-bus and serves as a Q-bus to DSSI adapter.

- Allows a DSSI bus with attached ISEs to be connected to a Q-bus
- Can be used with newer Q-bus systems as well as older MicroVAX-II systems
- Expands ISE storage capacity
- Passes requests and responses to and from the host
- Transfers data to and from the ISEs on the DSSI bus
- Supports dual-hosting of DSSI ISEs
- Does not support host-to-host cluster (SCS) communication across the DSSI bus

clusterverkeer

MASS STORAGE COMPONENTS

There are many components that provide mass storage in a VAXcluster system.

Disks and Disk Drives

- RA series disk drives
- RD series disk drives
- RF series Integrated Storage Elements (ISEs)
- RV series optical disks and drives
- RRD series CDROM disk drives

Controllers and I/O servers

- HSC series
- KDA/KDB/UDA series
- RQDX3 controller

Dual Path Configurations

Dual-pathed RA series disks support static dual access.

- Each dual-pathed RA disk is served by one controller (HSC, KDA, UDA, KDB) at a time. This is technically true of ISEs on the DSSI bus, for the controller and disk form one unit.
- If one of the paths or nodes locally connected to the disk fails, all other nodes automatically use the remaining, available path.

Dual-pathed MASSBUS and DSSI disks support dynamic dual access. Since the DSSI controller can accept service requests from multiple hosts simultaneously (like the HSC unit), it appears to behave like dynamic dual access.

- They permit simultaneous READ-WRITE access by all nodes connected locally to the disk.
- When the disk is MSCP served from both nodes, other nodes access the disk through either CPU path.
- If one of the nodes connected to the disk loses access (failed path), all nodes not directly connected to the disk (including the node whose path failed) automatically use the available path.

HSC Disk Performance

The HSC rarely becomes overloaded

- Individual disks generally become overloaded first.
- Use of multiple channels can improve performance if channels become overloaded.

The HSC optimizes I/O by:

- Performing I/O operations in the most efficient order, not necessarily the order in which they were requested
- Breaking an I/O request into fragments, if necessary, for efficiency

MSCP Served Disk Performance

The throughput of a MSCP served disk is very close to that of a local disk.

- Sufficient memory must be allocated for the MSCP server buffers.
- AUTOGEN allocates buffer space according to MSCP server workload.

Tapes and Tape Drives

The following tape drives are available:

- TA and TU series
- TS series
- TK series

Issues that Affect Disk and Tape Selection

- Application and operational needs
- The type of controllers used
- The type of cluster - local area VAXcluster, CI, DSSI, mixed-interconnect
- The number of disk and tape drives needed for effective storage capacity
- Availability requirements for the cluster

VAXcluster CONSOLE SYSTEM

optioneel produkt
Geen cluster nodes
tot 32 consoles
min. V5.3 van VMS

Large clusters have many separate console devices:

- VAX processors
- MicroVAX processors and VAXstation processors
- HSC units

The VAXcluster Console System

- Is not a member of the cluster
- Is a full-functioning, standalone VMS node
- Optional product
  - Allows console I/O to be handled by a single processor
  - MicroVAX or VAXstation processor
- Can be connected to all VAX nodes
  - Use fiber-optic cable
  - Connect directly to the console port
- Can be connected to MicroVAX and VAXstation members distributed throughout the Ethernet
  - Use a terminal server connection to the member's console port.
- Each member's console log is individually accessible by a multiwindow display, video terminal, or VAXstation screen.
- Console commands can be given to any node in the cluster from the single VAXcluster console system.
- This product is based on network communications and is not limited to VAXcluster environments.

PREPACKAGED VAXcluster SYSTEMS

It is possible to purchase some VAXcluster system configurations with a single model number.

Table 2–1 Examples of Prepackaged VAXcluster System Attributes

	VAX 6000-312	VAX 6000-423	VAX 6000-424
Processor	2 VAX 6000-310	3 VAX 6000-420	4 VAX 6000-420
Total Memory	64 Mbytes	192 Mbytes	256 Mbytes
Storage Controller	1 HSC40	1 HSC70	1 HSC70
Disk	1 SA600 (4 RA90)	8 SA600 (64 RA90)	12 SA600 (96 RA90)
Tape	1 TA79	2 TA90	2 TA90
Console	VT330	VCS	VCS
Terminal Server	1 DECserver 200	1 DECserver 200	1 DECserver 200
Ethernet	2 DEBNA, 1 DELNI	3 DEBNA, 1 DELNI	4 DEBNA, 1 DELNI

- VAX 6000-312 dual-processor, CI based VAXcluster system, based on VAX 6000-310
- VAX 6000-423 triple-processor, CI based VAXcluster system, based on VAX 6000-420
- VAX 6000-424 quad-processor, CI based VAXcluster system, based on VAX 6000-420

Processor Configuration Issues

- Advantages and disadvantages of each type of interconnect
- Available hardware
  - To best make use of interconnect and communication hardware you already have
- The types and locations of CPUs
  - VAXcluster systems using CI and DSSI interconnects are limited by cable length.
  - DSSI bus can be directly connected to Q-bus systems.
- Present and future expansion needs
  - There are smaller limits to the number of nodes supported in clusters using CI and DSSI interconnects.
  - See the VAXcluster SPD for current restrictions.
- Cost of interconnect and installation
- The availability requirements of your cluster
  - The CI cluster is the most highly available system.
  - Clusters with dual-hosted DSSI disks are also highly available.
  - The mixed-interconnect cluster can be configured to be as highly available as the CI cluster, but the Ethernet is still a single point of failure.
  - The local area VAXcluster system can be configured with dual boot servers, dual-ported disks, and a quorum disk to increase availability.
- Expected load on the interconnect
  - The I/O and computation load in a CI cluster should be spread as evenly as possible among CI members to avoid overloading.

SUMMARY

The hardware components of VAXcluster systems include:

- VAX processors
- The three interconnects:
  - Ethernet
  - CI bus
  - DSSI bus
- Mass storage components available for cluster storage needs
- The VAXcluster Console System
- Prepackaged VAXcluster systems available for purchase

APPENDIX — BOOT SERVER CAPACITY

Table 2–2 serves as a reference for configuration of mixed-interconnect and local area VAXcluster systems, since the I/O capacity of the boot server processors is very important.

Table 2–2 I/O Capacities of Representative Boot Servers

CPU Type	Adapter Type	I/Os per Second	Limiting Resource
VAX–11/750	DEUNA or DELUA	45	CPU or DEUNA
VAX 8300	DEBNA	50	CPU
VAX 8200	DEBNA	55	CPU
VAX 8350	DEBNA	60	CPU
VAX 8250	DEBNA	65	CPU
VAX–11/780	DELUA	70	CPU
VAX–11/785, VAX 8600 VAX 8650	DELUA	100	DELUA
VAX 8500, VAX 8530 VAX 8550, VAX 8700 VAX 8800, VAX 6000 series	DEBNA or DEBNI	135 (DEBNA) 340 (DEBNI)	DEBNI (for VAX 8800 series) DEBNA (for all others)
MicroVAX 2000	DESVA	20	ST506 Controller
MicroVAX II and RD Disks	DELQA	45	RQDX3 Controller
MicroVAX 3500	DELQA	60	2 RA70s
MicroVAX II and RA Disks	DELQA	80	CPU
MicroVAX 3600	DELQA	120	DELQA, 4 RA82s
MicroVAX 3300 MicroVAX 3400	On CPU module	130	CPU
MicroVAX 3800 MicroVAX 3900	DELQA DESQA	165	CPU

(MicroVAX 3300 and MicroVAX 3400 information in this table is with six RF30 disks.)

VAXcluster Software Concepts

INTRODUCTION

VAXcluster system integrity is created and maintained by the VAXcluster software. Upon this foundation of cluster membership and integrity, resources can be managed and made available cluster-wide.

Knowlege of the VAXcluster software components and their functions is useful for cluster management and troubleshooting.

OBJECTIVES

After completing this module, students should understand the concepts of:

- VAXcluster Communication Protocols
  - System Communication Services (SCS)
  - MSCP (Mass Storage Control Protocol)
- System-Supplied Components
  - Connection manager
  - Distributed lock manager
  - MSCP server
  - Distributed job controller
  - Cluster-wide operator communication (OPCOM)

RESOURCES

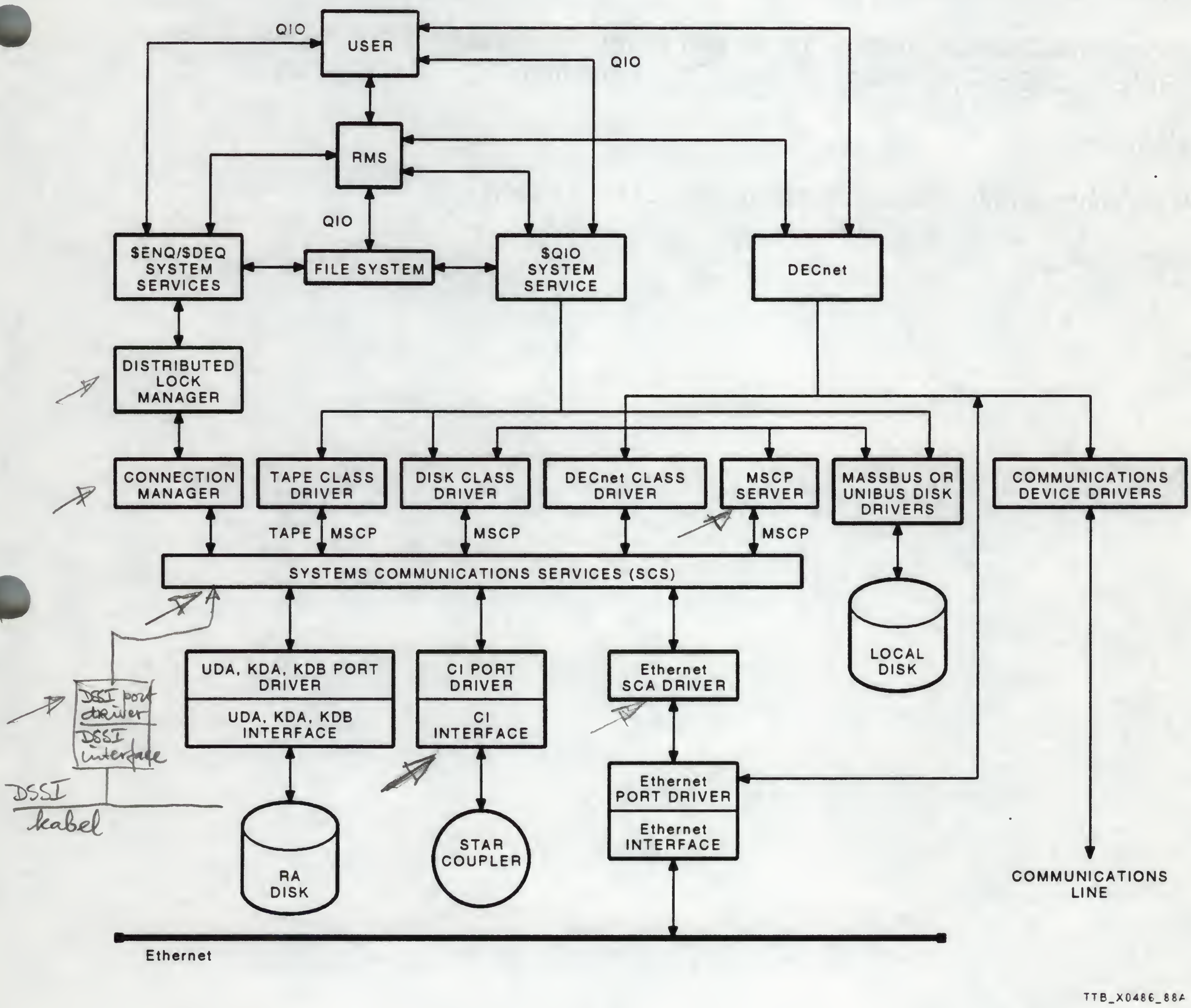
- *Networks and Communications Buyer's Guide*
- *VMS VAXcluster Manual*
- *VAX Systems/DECsystems Systems and Options Catalog*
- *VMS System Services Reference Manual*
- *VMS Device Support Manual*
- *VAXcluster System and Application Design*

VAXcluster SOFTWARE

Several components of the VMS operating system are designed specifically for coordination and communication within a VAXcluster system. These components are closely integrated with the operating system. They are installed automatically when you install the VMS operating system. However, you are not permitted to use them without a valid software license. Some knowledge of the VAXcluster software functions is useful for cluster management and troubleshooting.

Figure 3–1 discusses the various software components of a VAXcluster system.

Figure 3–1 VAXcluster Software Overview



in cluster

VAXcluster COMMUNICATION PROTOCOLS

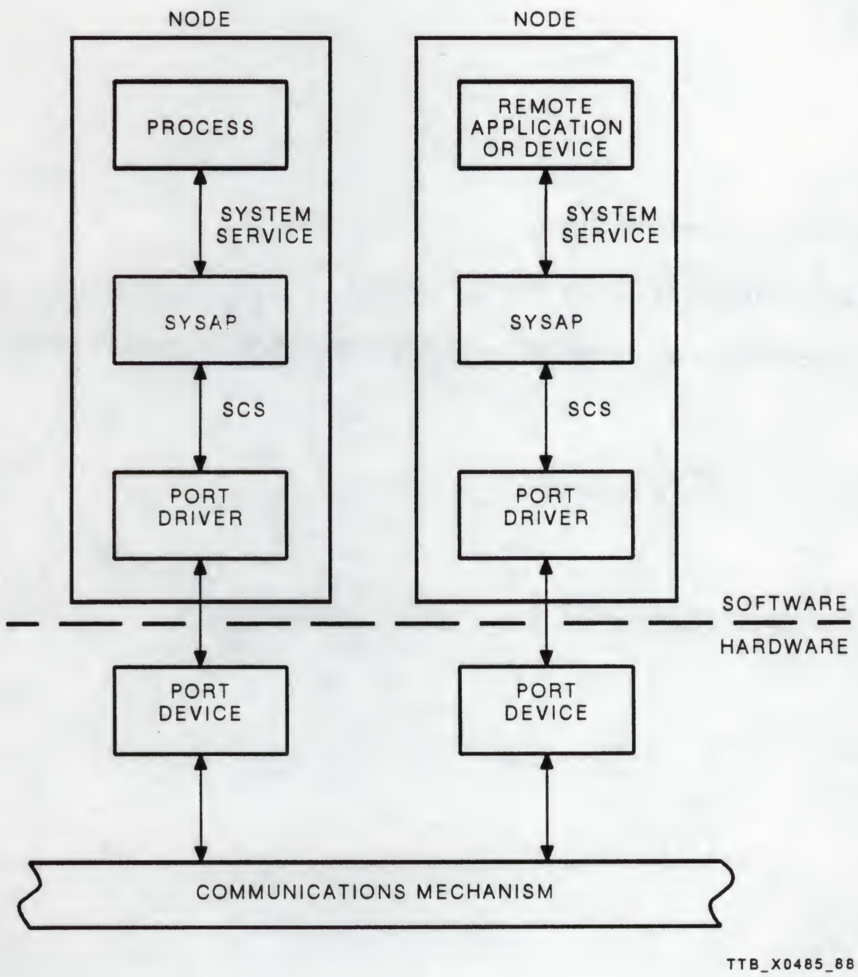
The framework for VAXcluster communication is provided by the System Communication Architecture (SCA). This layered architecture provides for connections made between peer System Applications (SYSAPS) to communicate using System Communications Services (SCS) to talk through port drivers which send the messages out across a specific port to the appropriate interconnect.

System Communication Services (SCS)

SCS is used by VMS software to communicate with corresponding software on other nodes within the cluster. These communication services support the SYSAPS in a comprehensive System Communication Architecture (SCA).

- SCS communications go through the CI port driver (PADRIVER), Ethernet port driver (PEDRIVER), or DSSI port driver (PUDRIVER) to and from:
  - Members
  - HSC controllers (which communicate over CI buses only)
  - DSSI devices

Figure 3–2 How System Service Calls Use SYSAPS to Communicate Across the Cluster



Each SYSAP on one VAXcluster node must communicate with a corresponding SYSAP on another node, for example:

- Connection manager must communicate with connection manager
- Disk class driver must communicate with MSCP server

SCS routines:

- Implement Digital's Systems Communication Architecture (SCA)
  - SYSAPs
  - SCS
  - Port Drivers
  - Ports
- Establish and break connections between members
- Format and transfer messages between SYSAPs and port drivers
- Do not perform routing of messages — all members must have direct connections to each other

Class and Port Drivers

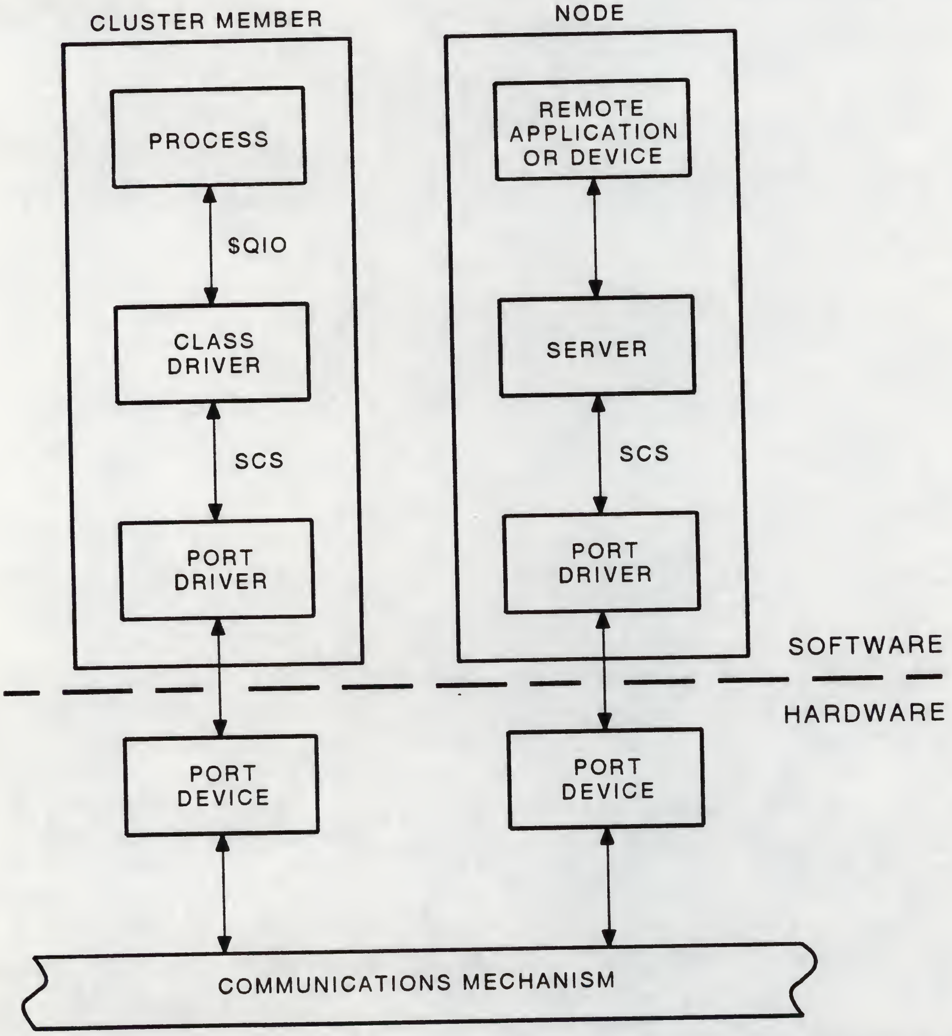
Device drivers handle communication between user processes and local physical devices.

A class driver communicates with user processes and handles one class of device:

- Disk class driver (DUDRIVER) for disk I/O traffic
- Tape class driver (TUDRIVER) for tape I/O traffic

Figure 3–3 shows the relationship among class drivers, port drivers, and Systems Communications Services (SCS).

Figure 3–3 Class Drivers, Port Drivers, and SCS



TTB_X0487_88A

MSCP (Mass Storage Control Protocol) Server

The MSCP (Mass Storage Control Protocol) server is used for cluster communication between a host system and a device controller.

Disk controllers that use the MSCP server include:

- UDA, KDA, KDB, KDM, and RQDX series of controllers
- HSC controllers
- ISE disks
- Any VMS cluster member, when running the MSCP server

I/O requests to MSCP controllers go through a class driver. When a process requests a QIO to a disk on an MSCP controller:

- A request goes to the appropriate class driver and is:
  - Converted to MSCP by the class driver
  - Formatted and sent to the port driver by SCS
  - Decoded by the disk controller on the remote node
- After the request is decoded:
  - The controller transfers data in blocks through the port driver.
  - The controller sends MSCP messages through SCS to the originating class driver, which notifies the requesting process.

The Tape Mass Storage Control Protocol (TMSCP) is used similarly for tape controllers:

- TQK50 controller
- HSC series controllers (TA tapes)

VAXcluster SYSTEM-SUPPLIED COMPONENTS

These system-supplied components are present on each VMS cluster member:

- Connection manager
- Distributed lock manager
- Distributed transaction manager
- MSCP server
- Distributed job controller
- Distributed Operator Communications (OPCOM)

Each of these components contributes to the concept and "feel" of the cluster environment from a different perspective:

- Cooperating VMS systems
- Cluster resources
- Disk "localness"
- Queues
- Operator information

The following utilities assist in the management of a VAXcluster system and are discussed in subsequent modules:

- System Management Utility (SYSMAN)
- License Management Facility (LMF)
- Network Control Program (NCP) and DECnet software
- Local Area Transport Control Program

Connection Manager

Connection manager software runs on each active member in a cluster to coordinate the state of the cluster as members join or leave.

The connection manager:

- Determines what nodes are in the cluster
- Allows each active node to be aware of cluster membership
- Conducts cluster state changes
- Provides the necessary membership support to the lock manager
- Constructs and maintains the data structures needed to ensure a uniform view of the cluster from each node
- Uses SCS to deliver messages to other nodes for upper levels of software
- Uses a quorum scheme to prevent cluster **partitioning**.
  - Partitioning occurs when two or more clusters have access to common resources and are unaware of each other.
  - Partitioning can cause disk file corruption, since there must be coordination among all cluster members sharing a file system.

The Quorum Scheme

Cluster Quorum is a dynamic value calculated by the connection manager to prevent partitioning. It allows processing to occur only if a majority of the expected voting member nodes are functioning.

Each cluster member is assigned a fixed number of votes that it contributes toward quorum:

- Satellites should be given zero votes (default).
- The default for each non-satellite member is one.

During a **cluster state transition,** the connection manager totals the number of votes of all members present, and compares that value to the **cluster quorum value.**

- The cluster runs if the total number of votes is at least the cluster quorum value,
- If the cluster votes value is less than the cluster quorum value, the cluster suspends processing until enough votes are present.
- Cluster state transitions occur when a node joins or leaves the cluster, and when the cluster recognizes a **quorum disk.**

The quorum disk acts as a virtual node, and is given votes to add to the cluster votes total.

- It increases the availability of small configurations (notably two-member clusters).
- For a quorum disk to be used:
  - One or more nodes must have a direct connection to the disk.
    These nodes are called **quorum disk watchers.**
  - The SYSGEN parameter DISK_QUORUM on each quorum disk watcher is set to the disk's device name. It is left blank for the other nodes in the cluster.
  - The remaining nodes recognize the specified name by communicating with the quorum disk watcher.

Cluster Quorum Value and the EXPECTED_VOTES Parameter

The cluster quorum value is initially set by using the value of the SYSGEN parameter EXPECTED_VOTES to calculate the minimum number of votes necessary to ensure that partitioning does not occur.

- The system manager should set EXPECTED_VOTES to the sum of all votes held by potential cluster members (and quorum disk, if any).
- The initial quorum value is calculated using the following formula:

```
INTEGER((EXPECTED_VOTES + 2) / 2)
```

When a cluster state transition occurs, the connection manager recalculates quorum using the maximum of:

- The current cluster quorum value
- The integer value of (EXPECTED_VOTES + 2) / 2 on the booting node
- The integer value of (V + 2) / 2, where V is the total number of votes held by present members

A VAX system is not allowed to join the cluster if it specifies an EXPECTED_VOTES value that would cause the cluster to suspend activity.

The cluster quorum value is not reduced automatically when a VAX system leaves the cluster.

cluster exp votes is het max

- Exp votes
- V tot
- current cluster exp votes

Distributed Lock Manager

The distributed lock manager is a tool used by the operating system and available to programmers.

Locks are software records keeping track of resource usage. They are used to coordinate access to and prevent conflicting use of those resources.

- Used to synchronize access to shared cluster-wide resources:
  - Devices
  - Files
  - Records in files
  - Any user-defined nameable resource
- Used for cluster-wide communication through the lock value block
- Releases locks when a node fails, so that the remaining nodes may continue processing

The distributed lock manager is used by:

- The file system (XQP)
- VAX RMS (Record Management Services)
- The distributed job controller
- User-written VAXcluster applications

The distributed lock manager uses the connection manager and SCS to communicate information among members of the cluster.

- Each resource is managed cluster-wide by the first node to lock the resource (the resource master).
- Each resource has a directory node that knows the resource master.
  - Nodes with LOCKDIRWT $\neq$ 0 participate in the directory function. (ongelijk ∅)
  - If all nodes have the same LOCKDIRWT, including 0, they participate equally.
- Mastery is subject to change depending on the nodes that lock the resource.

LOCKDIRWT indicates an "eagerness" to manage locks. Lock trees can be moved from one node to another when a number of nodes are locking in the same resource tree and one of those nodes is more "eager" to manage locks. This will tend, over time, to make lock management more efficient and to reduce lock rebuild times.

The distribution procedure operates as follows:

- If multiple nodes with a non-zero value for LOCKDIRWT have locks on a resource, the directory node is favored as the resource manager, if it holds locks on the resource.
- If a node with a zero value for LOCKDIRWT and a node with a non-zero value have locks on the same resource, then the non-zero node is favored as the resource manager.
- Resources locked by processes on only one node continue to be managed by that node.

During a cluster-state transition, the lock database is reconstituted as necessary.

- If a node with a zero value for LOCKDIRWT joins a cluster containing at least one node with nonzero LOCKDIRWT, only a merge rebuild is done.
- If a node with a zero value for LOCKDIRWT is leaving a cluster containing at least one node with nonzero LOCKDIRWT, only a partial rebuild is done. This results in any locks held by the departing node being released, and any trees it is mastering being remastered to other nodes using those trees.
- If a node with nonzero LOCKDIRWT is leaving or joining a cluster, a "directory rebuild" is done:
  - All directory node information is discarded throughout the cluster, and the Lock Directory Weight Vector is rebuilt.
  - As a result of this table being rebuilt, directory responsibility for some resource trees may shift. Thus, resource managers inform directory nodes where all trees are now being mastered.
  - When such a node leaves the cluster, a partial rebuild is also required to remaster any trees the departed node was mastering which include locks from other nodes.
- If the node that fails has LOCKDIRWT set to zero, and there is a single node in the cluster with LOCKDIRWT of one or greater, then a "fast rebuild" is done.
- You may wish to set LOCKDIRWT to zero on MicroVAX and workstation members.

Cluster-Wide Access to Files

The Distributed Lock Manager permits coordinated use of resources throughout the VAXcluster system. This allows users of the cooperating systems access to the files associated with cluster-available devices on any processor or intelligent, MSCP-compliant controller:

- Devices connected to an HSC controller or DSSI bus
- Local devices that are served to the cluster using the MSCP server

VAX RMS disk files can be shared cluster-wide to the record level.

MSCP Server

The MSCP server allows disks connected locally to a VAX processor or to an intelligent, MSCP compliant controller to be shared cluster-wide.

- These disks include:
  - Disks local to CI members
  - Disks on boot servers
  - Disks on disk servers anywhere in the cluster, including satellites
  - HSC disks in mixed-interconnect clusters (when the MSCP server is running on one of the CI nodes)
  - ISE disks in a mixed-interconnect cluster
- The MSCP server decodes and services MSCP I/O requests sent by the disk class driver on remote cluster nodes.
  - Once a device is MSCP served, any processor in the cluster can mount the device and access it.
- The MSCP server is controlled by answers to questions in CLUSTER_CONFIG.COM, which sets the SYSGEN parameters MSCP_LOAD and MSCP_SERVE_ALL.
  - MSCP server is turned off by default on Ethernet satellites.
  - MSCP server is turned on by default on boot server and disk server nodes.

Distributed Job Controller

The job controller distributes the batch and print processing workload over cluster nodes:

- Permits users to submit batch and print jobs to queues that execute on any node in the cluster
- sys$common
  Uses a common queue file, JBCSYSQUE.DAT, to maintain the current state of all queues on all systems on the cluster
- Allows generic queues to be created that feed equivalent specific queues on any systems in the cluster → t.b.v. workload verdeling
- Directs batch and print jobs to the execution queue with the lowest ratio of jobs to queue limit (or to the next available queue)
- Uses the distributed lock manager to signal other VAX nodes to examine the batch and print queues for jobs to be processed

Distributed Operator Communications (OPCOM)

The Distributed Operator Communications (OPCOM) allows operator terminals to operate cluster-wide.

- Any cluster member can receive, record, and respond to operator messages sent from other members.
- Terminals on any node can be directly addressed by OPCOM messages.
- All enabled terminals in the cluster receive all OPCOM messages in the cluster.
- Each OPERATOR.LOG contains messages from all nodes.

SUMMARY

The features of VAXcluster software include:

- System Communication Architecture (SCA)
- SCS protocols for cluster communication
  - Class and port drivers that use SCS
- MSCP protocols for data storage
- VAXcluster system component software:
  - Connection manager
  - Distributed lock manager
  - Distributed transaction manager
  - MSCP server
  - Distributed job controller
  - OPCOM

APPENDIX — OPTIONAL PRODUCTS

Products that may be useful in a cluster:

- VAX RMS Journaling
- Distributed System Services
  - Remote System Management (RSM)
  - Distributed Queuing Service (DQS)
  - Distributed File Service (DFS)

VAX RMS Journaling

VAX RMS Journaling is an optional product that offers a set of file protection mechanisms that can be set up on any VAX RMS file, protecting it against three different types of failures:

Hardware Failures

- Hardware failures can cause data that has been written to be lost.
- After-Image (AI) journaling copies all file modifications of a file to a journal file. If the original file becomes unusable, it can be recovered using the journal file.

Software Errors

- Software errors might delete data in a file.
- Before-Image (BI) journaling copies the existing data into a journal file before a write is done to a file. If any of the old data is needed, it can be recovered from the journal file.

Incomplete Operation Sets

- It is possible that a set of writes needs to finish completely, or not at all. If a hardware or software problem prevents one of the individual disk operations, then none should be written.
- Recovery-Unit (RU) journaling allows a set of operations, called a **recovery unit**, to complete before any of the operations are written to disk.

Distributed System Services

Remote System Management (RSM)

RSM allows network nodes to be managed from another network node.

- The operating system is down-line loaded and all management operations, such as backup and changing user environment characteristics, are done remotely.
- RSM is an alternative for local area configurations where centralized management of many MicroVAX and other systems is desired, but a local area cluster is not appropriate.
- RSM can also be used to manage multiple local area VAXcluster systems on one LAN from a single centralized point.

Planning a VAXcluster System

Distributed Queuing Service (DQS)

DQS allows network nodes to submit print jobs to a network print server:

- Allows many different forms and printers
- Is an alternative to configuring a print server member in each cluster, for network configurations including several cluster members and standalone nodes
- Is accessible from multiple operating systems (TOPS-20, TOPS-10, RSX, ULTRIX)

Distributed File Service (DFS)

DFS allows disks to be served across a network:

- Disks can be local to a specific node, reside on an HSC, or a DSSI bus.
- The member that provides DFS services is a DFS **server system.**
- Remote systems that mount these disks are called DFS **client nodes.**
- Any user on any node can perform most operations on the DFS-served disk as if the disk were mounted locally, except for shared write access to VAX RMS files.

INTRODUCTION

A VAXcluster environment may be configured in many ways to address different application and user needs. To obtain the configuration that is best for your application, you must spend time planning. This module presents information that you will use in preparation for building and managing your VAXcluster environment.

OBJECTIVES

After completing this module, students should be able to:

- Plan cluster security and management
- Plan cluster mass storage
- Determine VAXcluster configurations best suited for a given workload
- Determine the proper network configuration for a given VAXcluster environment
- Determine the proper set up of the hardware
- Assign a name to each device in a cluster
- Select appropriate values for the cluster parameters

RESOURCES

- *VMS VAXcluster Manual*
- *VMS Networking Manual*
- *Guide to Setting Up a VMS System*
- *Guide to VMS Performance Management*
- *Guidelines for VAXcluster Configuration*
- *Introduction to VAXcluster Application Deisgn*

VAXcluster MANAGEMENT AND SECURITY

Management Considerations

A VAXcluster system is a single management domain. It must be managed as a single system, by a single system manager, or a cooperating management team. Even if a cluster has many different working environments, the different systems have access to common, shared resources and must be guided by a single security and management policy.

Members of a system management team must decide how to share the following:

- Troubleshooting duties
- Maintenance duties
  - VMS updates and upgrades
  - Security monitoring
- System expenses
- System resources
  - Mass storage devices
  - Printers
  - Batch queues

Security Considerations

- Privileged users on one VAX node can affect the other nodes.
- There is no node-specific file protection by default.
  - — The system manager can implement node-specific protection
  - — Use a coordinated rights list and Access Control Lists (ACLs)
- A network can provide greater security isolation between nodes than a VAXcluster environment.
- A single system implies that user names, UICs, and access rights are unique throughout the cluster.
- Multiple User Authorization Files
  - — Not recommended in most cases
  - — Break the single-system model of a cluster
  - — Must be kept consistent manually
  - — Do not isolate privileged users

MASS STORAGE CONFIGURATIONS

Consider the following in planning an effective mass storage configuration:

- Disk configuration
- System disk configuration

Disk Configuration Considerations

Methods of sharing disks among all cluster members:

- Connect the disk to an HSC or DSSI unit
- Serve the disk to other members of the cluster over the Ethernet. (NI - cluster)
- Connect the disk locally to a particular member to be MSCP served. MSCP_LOAD = 1
  MSCP_SERVE_ALL = 1 or 2
  or $SET DEVICE/SERVED dev:
- Provide dual paths for MSCP served disk.
  - Failover and load balancing are automatic.
  - Increases the availability of HSC and DSSI disks in mixed-interconnect and local area clusters.
  - Increases the availability of local controller disks in any cluster configuration where such disks are used.
  - The device name must be unique and independent of the path used.

Methods of restricting disk access to users on a particular VAX node:

- Connect the disk locally
- Do not MSCP serve the disk
- MSCP serve the disk but do not mount it
- MSCP serve the disk and mount it on selected members
- Use ACLs to restrict access

System Disk Configuration

A VAXcluster system disk configuration can include:

- One or more common system disks
  - Contains the operating system for one or more members
  - All systems booting from the disk share most operating system files
- See Appendix A of Module 5 for more detailed information about System Disk structure.

A system disk can be:

- Connected to a single HSC unit
- Dual-ported between a pair of HSC units
- Hosted by a MicroVAX on the DSSI bus
- Dual-hosted between two MicroVAX systems on a DSSI bus
- Connected locally to any cluster member
- Dual-ported between cluster members that have separate, single ported system disks for booting
  - Systems cannot boot from a local dual-ported disk.
  - Satellites can boot remotely from dual-ported disks through the locally connected boot servers.

Advantages of a Common System Disk

Management risks are reduced by the use of one or more common system disks.

The following operations need be performed only once for each system disk:

- VMS installations
- VMS upgrades
- VMS updates
- Layered product installations
- System disk backups

There are limits on systems sharing a common system disk.

- Tradeoff is ease of management versus performance.
- SYSDUMP.DMP files of large memory systems can take up significant amounts of disk space, determining the maximum number of systems sharing a system disk.
- Partial dumps save significant amounts of disk space.
- SYSDUMP.DMP is not required, if space is at a premium.
- A common SYSDUMP.DMP can be established, with pointers established using:

```
$  SET FILE/ENTER=SYS$SYSDEVICE:[SYSn.SYSEXE]SYSDUMP.DMP -
_$ SYS$COMMON:[SYSEXE]SYSDUMP.DMP
```

System Disk Availability

All systems booting from a system disk fail when the system disk or its controller fails.

Possible solutions for this problem are:

- Several common system disks
- Shadowing the system disk
- Dual-ported or dual-pathed system disks

System Disk Performance

A heavily used system disk can be a bottleneck that degrades performance of the entire VAXcluster system.

- The number of systems you can run without degrading performance depends on the total load on the system disk.
- To alleviate a system-disk bottleneck:
  - — Reduce the load on the disk by moving some system files to other cluster-available disks.
  - — Volume shadowing may ease the bottleneck (for read I/O), as well as eliminating a single point of failure.

PERFORMANCE CONSIDERATIONS

pagen en swappen bij voorkeur naar niet-system disken

Effect of I/O Workload on Performance

- A VAXcluster system has greater potential for overloading disks than does a single system.
  - Multiple systems can make I/O requests to the same disk.
  - A disk is often the primary performance bottleneck.
- The greater the number of users who share a resource, the more synchronization is needed.
- Directing the paging and swapping activity to another disk will reduce the load on the system disk.
- Configuration recommendations depend on workload.
  - If there is no sharing at file level (for example, each user has his own files), balance the load among available disks, or add disks.
  - If there is sharing at file level, tune the system and applications to decrease disk I/O and synchronization.

I/O Performance Using Ethernet

Bij satelliet geen "quorum-disk" opgeven
Eventueel een bridge ertussen

The Ethernet rarely overloads. Usually the performance problems can be traced to CPU overloading in older/slower CPUs or to Ethernet controller overloads with newer/faster CPUs.

- If the Ethernet I/O rate exceeds 200-300 I/Os per second, overload may occur.
- If total I/O from any individual member exceeds 70-115 I/Os per second, overload may occur.
- If there is an overload:
  - Consider reconfigurations that remove some load from the Ethernet.
  - Consider isolating groups of the cluster into several smaller clusters using LAN bridges.

Satellite I/O to boot server disks

- Takes about 25 percent longer to complete.

I/O Performance Using the CI Bus

- Spread I/O as evenly as possible among members.

New Ways to Improve Disk Performance

- Striping
- ESE20

Contention for Resources

Contention for resources can affect processor performance within a cluster, depending on the interconnect chosen.

- On MicroVAX DSSI systems:
  - Each ISE disk has its own controller.
- In an Ethernet or mixed-interconnect VAXcluster system:
  - Boot servers must be chosen with adequate power and fast Ethernet controller.
  - "Fast" VAX processors have to wait for "slow" processors to finish using a resource.

The Distributed Lock Manager:

- Lessens contention by synchronizing access to shared resources.
- Overhead is usually insignificant on a well-tuned system.
- That overhead is **independent** of the number of nodes in the VAXcluster system.
- When associated with an I/O operation, overhead is a fraction of the total CPU time taken by the operation.
- Clusters usually become I/O-bound before locking overhead becomes significant.

Memory Requirements

- VAXcluster members require at least four megabytes of physical memory.
- VAXcluster systems running applications locally need more memory.
- When increasing the number of nodes, an increase in memory can be beneficial.
- Satellites should be configured with enough memory to page infrequently and never swap.
- Large amounts of paging and swapping over the Ethernet can degrade performance in the cluster and on the network.

CPU Workload

A CPU-bound application can take full advantage of added VAX processors, if the load is balanced among the CPUs in the cluster.

Adding CPUs to a cluster:

- Can turn a CPU-bound application into an I/O-bound application

CPU load balancing

- The job controller allows balancing of the batch job load when a generic batch queue is associated with more than one execution queue:
  - When a job on the generic queue is scheduled for execution, the job controller assigns it to the execution queue with the lowest ratio of active jobs to the job limit.
  - If execution queues are equally loaded, the job controller assigns the next job to the first execution queue assigned to the generic queue.
- Terminal servers balance the interactive user load.
  - Each VAX system computes a **service rating** based on CPU type, percentage of idle time over a recent interval, and number of interactive jobs versus interactive job limit.
  - The terminal server connects each user to the VAXcluster member with the highest service rating if a common service is provided.
  - Service rating is usually, but not always, a good tool for load balancing.
  - Terminal servers allow user selection as well. Users tend to choose the nodes that have the best response time.
  - Terminal server load balancing is static. A user, once logged in, remains attached to the same node.

Single-Node performance

- A single overloaded and/or poorly tuned system can degrade performance of the entire VAXcluster system.
- If tuning becomes necessary, tune individual nodes first.

Queue Environment

All queues can be made accessible to all nodes in a cluster.

- ACLs can be used to restrict the use of a queue to specific nodes or users.
- Batch and print jobs can be submitted to:
  - Execution or logical queues for processing on a particular VMS system or printer
  - Generic queues, which use a load-balancing scheme to direct processing to an execution queue on any active node
- If a node fails, jobs submitted to generic batch queues can make use of the batch restart capability to run on another node.

Terminal Configuration

Connecting a terminal server and all cluster nodes to the same Ethernet:

- Allows access to any of the nodes from one terminal
- Increases CPU availability
- Allows the creation of simultaneous sessions on any node(s) in the cluster
- Can be used to balance the interactive user load among nodes providing the same LAT service

SELECTION OF SPECIAL PURPOSE MEMBERS

Boot Servers

General issues involved in boot server selection:

- High CPU performance
- High-performance Ethernet controllers bv. DELQUNA i.p.v. DEUNA
- Sufficient disk capacity on the system disk niet > 80% bezetting
- Sufficient I/O throughput
- Application-dependent boot server needs (queues, printers, tapes)
- Physical location of processors
  - — The boot server node closest to a satellite services a down-line load request.
  - — If possible, make sure the closest node is not one of your slower processors.

NETWORK CONSIDERATIONS

Although the VMS operating system does not use DECnet software for VAXcluster communication, Digital requires you to install and run the DECnet VAX product in your VAXcluster system.

for Down-line loading.

DECnet Software in a VAXcluster Environment

All VAXcluster CPUs require either an end node or full function DECnet VAX license.

- Makes the VAXcluster environment easier to manage
- Allows access to disks that are not available cluster-wide
- Allows login on any VAXcluster node from any terminal using the SET HOST command
- Allows distributed applications that require DECnet task-to-task communication to work in a cluster
- Allows down-line loading of satellites and other Ethernet devices in Ethernet configurations
- Allows the VAXcluster system to exist within a larger, networked environment
- Allows the use of the cluster-wide functions of the MONITOR utility and other system management tools

Using the CI for Network Communication

- Eliminates the need for Ethernet in CI only configurations
- Allows DECnet operations on CI nodes
- Requires a DECnet routing node if there are three members or more in the cluster. Two routing nodes ensure DECnet network availability if one router fails.

Using Ethernet for Network Communication

- Yields high network throughput and less overhead than the CI DECnet
- Allows VAXcluster nodes to be part of a larger network
- Is the same medium used for other network servers
- Is used for down-line loading of workstations and MicroVAX nodes booting as satellites into the cluster

PREPARING TO BUILD A VAXcluster

Device Names in a VAXcluster System

Guidelines:

- Each single-ported device must have a unique device name within the cluster.
- To change the device name you must change the unit number (unless you are actually changing the allocation class or SCSNODE of the host, or similar parameters on a DSSI or HSC device).
- Each dual-ported device must have the same physical device name ($allocation_class$device_name) on each node connected to the device.
- DSSI Integrated Storage Elements and HSC disks must have the same allocation class as the VAX systems that MSCP serve them to the cluster across the Ethernet.

Two formats are used in naming devices:

- **node$device**
  - — Used for devices that are directly connected to only one node. In Figure 4–1, LION$DUA0 is named using this format.
  - — Zero is the default value for the ALLOCLASS SYSGEN parameter, forcing this format.
- **$allocation-class$device**

  op satelliet altijd ∅

  - — Setting the ALLOCLASS parameter to non-zero (1-255) enables this format.
  - — Used for dual-pathed (including dual-ported and dual-hosted) devices. These cannot have zero allocation class.
  - — Any disk MSCP served to the cluster must have the same allocation class as the VMS member(s) that serves it to other members.

When a device is on a local node:

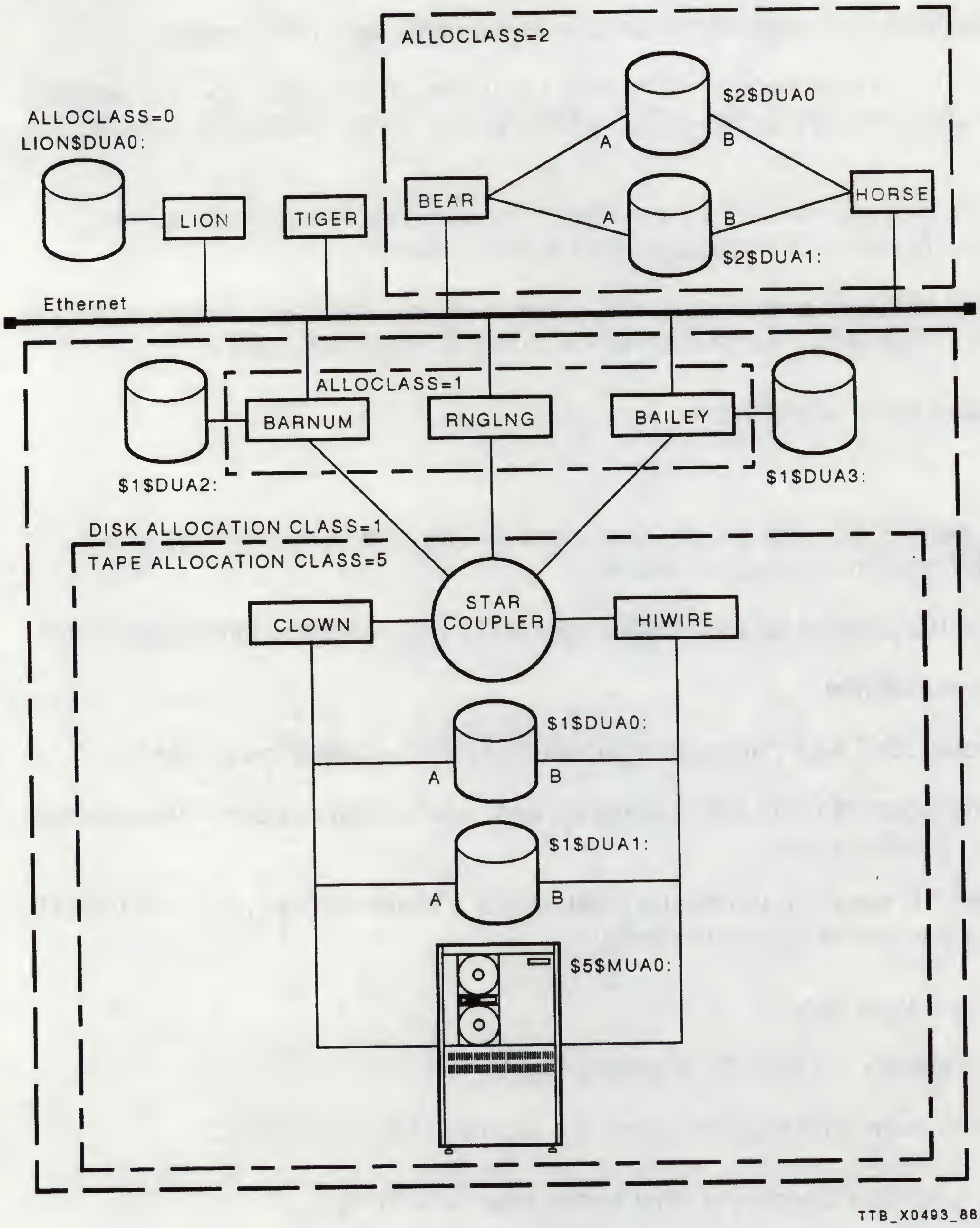
- You can use its traditional name (for example, TXA2).
- You can prefix its name with the node name (for example, BARNUM$TXA2).

Use **node$device** to refer to devices on other nodes when specifying:

- A terminal on another node to OPCOM
- A printer on another node to the job controller

Figure 4–1 shows the sample cluster with allocation classes assigned to nodes. It also shows the device name of each disk and tape drive.

Figure 4–1 Disk and Tape Device Names in a VAXcluster System Using Allocation Classes



VAXcluster Parameters

Table 4–1 presents a checklist of questions that need to be answered before building a cluster. Each question corresponds to a SYSGEN or network parameter that must be set when the cluster is built in order for the configuration to succeed.

Table 4–1 Information Requested for Installations and Adding Nodes

Question	Response	Parameter
Will the node be a cluster member?	Enter Y. The parameter will be set to 2. Other values for the parameter can be: 0 = Node will not participate in a cluster. 1 = Node will participate in a cluster if the hardware supporting SCS is present (CI, UDA, HSC). 2 = Node will always participate in a cluster.	VAXCLUSTER
What is the node's DECnet node name?	Consists of six or fewer alphanumeric characters, with at least one alphabetic character. • Should be unique within the network. • Will also be asked when NETCONFIG.COM is executed.	SCSNODE
What is the node's DECnet address?	Consists of an area number (from 1 to 63) and a node number within the area (from 1 to 1023). • Nodes in a cluster will probably be in the same area. • The SYSGEN parameter is calculated for you by multiplying the area number by 1024 and adding the node number. • This is asked when NETCONFIG.COM is executed.	SCSSYSTEMID

Table 4–1 Information Requested for Installations and Adding Nodes (Cont.)

Question	Response	Parameter
Will the node be a satellite? (or) Will Ethernet be used for cluster communications?	Parameter specifies whether the PEDRIVER is loaded to enable cluster communications over Ethernet. (0 = Don't load driver, 1 = Load driver) • If the node will be part of a CI only cluster, answer N. • If it will be part of a local area or mixed-interconnect cluster, answer Y.	NISCS_LOAD_PEA0
What is the device name of the system disk?	The default is the device from which the system is currently booted.	DECnet load assist parameter
What is the name of the new system root?	This is asked if you are adding a node to an existing cluster. • For a CI node, this is a value from 0-D (0 is the default when installing VMS). • For a satellite node, the value can be 10-FFFF.	DECnet load assist parameter
Will the node be a boot server?	Answer Y if this node will be responsible for down-line loading satellite boot files.	DECnet circuit characteristic: SERVICE ENABLED
Will the node be a disk server?	If this node is to serve its local disks and /or HSC disks, answer Y. • MSCP_LOAD and MSCP_SERVE_ ALL will be set to 1. • A boot server is automatically a disk server.	MSCP_LOAD MSCP_SERVE_ALL
Will the node serve HSC disks? RF series disks?	The HSC version is only asked if the node is a CI node. The RFxx version is asked of Q-bus processors on the Ethernet. If you answer no, MSCP_SERVE_ALL will be set to 2 (serve local disks only).	MSCP_SERVE_ALL

Table 4–1 Information Requested for Installations and Adding Nodes (Cont.)

Question	Response	Parameter
What is the value for ALLOCLASS?	If the system will serve HSC or DSSI disks, enter the HSC or ISE allocation class value. • If the system has dual-ported disks, enter a value from 1-255 that will be used on both sides. • Otherwise, enter 0.	ALLOCLASS
Does this cluster contain a quorum disk?	Enter Y if your configuration has a quorum disk.	
What is the name of the quorum disk?	You will be asked this question only if you answered yes to the above question. • There can be at most one quorum disk in the cluster. • Therefore, all nodes in the cluster that have access to the quorum disk should specify the same disk.	DISK_QUORUM
Where will the page and swap files be located?	The default is on the system disk. However, you can specify a local (non-HSC) disk.	
Do you want to allow conversational bootstraps? (Satellites only)	No is the default, for security. The parameter will be set to 0 to disable, 1 to enable.	NISCS_CONV_BOOT

Table 4–1 Information Requested for Installations and Adding Nodes (Cont.)

Question	Response	Parameter
What is the cluster group number?	Uniquely identifies each cluster present on the same Ethernet. Coordinates the assignment of this number with all other clusters on the same LAN. • Valid numbers are 1-4095 and 61440-65535 • Stored in SYS$COMMON:[SYSEXE]CLUSTER_AUTHORIZE.DAT	Bij Satelliet in SYS$SPECIFIC:[SYSEXE] cluster_authorize.dat
What is the cluster password? Onthouden of copiëren naar sys$specific	Can be from 1 to 31 alphanumeric characters in length and can include dollar signs and underscores. • Prevents clusters that somehow use the same cluster group number from interacting with each other's satellites • Stored in same file as cluster group number	

Table 4–1 Information Requested for Installations and Adding Nodes (Cont.)

Question	Response	Parameter

What is the Ethernet hardware address?	The address has the following form: xx-xx-xx-xx-xx-xx. You must include the dashes in your response. To obtain this value, DECnet VAX software must be running and Ethernet service enabled on the boot server.
	
	• For MicroVax II and VAXstation II satellites, enter the following commands at the satellite's console:
	
	`>>> B/100 XQ`
	`Bootfile: READ_ADDR`
	
	• For MicroVAX 2000 and VAXstation 2000 satellites, enter the following:
	
	`>>> T 53`
	`2 ?>>> 3`
	`>>> B/100 ES` of T 50
	`Bootfile: READ_ADDR`
	
	• For MicroVAX 3000 (4000) series satellites, enter the following:
	
	`>>> SHOW ETHERNET`
		DECnet node characteristic: HARDWARE ADDRESS

op elke vax !

>>> E/P/N !5/W 20001920

FF00 FF08

FF12 FF2B

FF56 FF34

op een draaiend systeem

$mc ncp

ncp> sho known line char

SUMMARY

To plan a VAXcluster environment, the cluster manager should take into consideration:

- Cluster security and management
- Cluster mass storage
- Processor configuration
- Special member selection
- Network configuration

Before building a VAXcluster environment, you should:

- Have Digital Customer Service bring the hardware to the correct revision levels.
- Make sure you know the name of each device in the VAXcluster environment and the Ethernet hardware address of each satellite.
- Decide the answers to the questions that will be asked during the VMS installation or CLUSTER_CONFIG.COM.

APPENDIX A — CREATING A CONFIGURATION TABLE

This is largely self-explanatory, and is included as a reference.

After you have determined what your device names and parameter settings will be, and before you begin building your VAXcluster environment, make tables that you can refer to during the building procedures. Tables 4–2 through 4–4 contain values for the sample cluster shown in Figure 4–1.

VAXcluster Parameters

Table 4–2 Parameters for the Sample VAXcluster Environment — Part 1

Node Name	CI Port Number	DECnet Address	Satellite?	System Disk Device
BARNUM	0	1.1	no	1DUA0
RNGLNG	1	1.2	no	1DUA0
BAILEY	2	1.3	no	1DUA0
LION		1.99	yes	1DUA0
TIGER		1.100	yes	1DUA0
BEAR		1.101	yes	1DUA0
HORSE		1.102	yes	1DUA0
HSC's { CLOWN	3			
HIWIRE	4			

nummers
op de
Star coupler

Table 4–3 Parameters for the Sample VAXcluster Environment — Part 2

Node Name	System Root	Boot Server?	Disk Server?	HSC Disk Server?	Disk Allocation Class
BARNUM	SYS0	yes	yes	yes	1
RNGLNG	SYS1	yes	yes	yes	1
BAILEY	SYS2	yes	yes	yes	1
LION	SYS10	no	no	no	0
TIGER	SYS11	no	no	no	0
BEAR	SYS12	no	yes	no	2
HORSE	SYS13	no	yes	no	2

Table 4–4 Parameters for the Sample VAXcluster Environment — Part 3

Node Name	Quorum Disk?	Page and Swap Device	Conversational Bootstrap?	Ethernet Hardware Address
BARNUM	no	1DUA0		
RNGLNG	no	1DUA0		
BAILEY	no	1DUA0		
LION	no	LION$DUA0	no	08-00-33-41-77-9F
TIGER	no	1DUA0	no	08-00-86-21-34-76
BEAR	no	2DUA0	no	08-00-44-09-03-70
HORSE	no	2DUA0	no	08-00-96-22-00-45

APPENDIX B — PLANNING CLUSTER APPLICATIONS

Efficient execution of application software is the primary objective of a VAXcluster system manager. Some applications are written with the VAXcluster environment in mind. Even if an application is not written to take advantage of VAXcluster features explicitly, it may work very well in a VAXcluster environment. When forming or expanding a cluster, it is important to consider how the target applications will interact with one another, running on the same member, or on different members.

Characteristics of effective cluster applications include:

- Shared cluster disks and files
- Shared queues and resources
- High availability
- Centralized application management
- Distributed use

Considerations Related to Cluster Applications

Applications that are especially suited to the VAXcluster environment include those that can process on different systems simultaneously (distributed applications) and can use shared data simultaneously. The advantage gained from running an application on a VAXcluster system depends on the number of processors the application can use simultaneously.

Examples of effective cluster applications include software development, information storage and retrieval, office automation, educational applications, and CAD/CAM and other distributed environments. In a software development environment, the computing load can be easily distributed by dividing users among the nodes of the cluster. Even with only one user, that user may edit a file on one node and compile a different file in batch on another node, making use of the computing power of the entire cluster.

There are applications that cannot easily be distributed. Consider a large simulation program in which each computation depends on the result of the previous computation, and the entire program takes many hours to run. If you run this program in a cluster, the CPU power of the other nodes is wasted (unless there are other applications to be run at the same time).

Remember:

- The cluster members pause when a system joins or leaves the cluster.
  - Under most conditions, these cluster state transitions are short (a few seconds to one minute) and are hardly noticeable to interactive users.
- A VAXcluster environment can be used effectively to process data collected in real time by other nodes to which it is connected in a network.
- A VAXcluster environment with some nodes dedicated to development and others to production may not be a good mix.
- A cluster can provide high availability, but not 100 percent fault tolerance.
- There is no automatic restart of processes from one member to another.
  - Applications can be written to provide automatic restart.

The same considerations for installing applications on a standalone VMS system hold for members of a cluster also. Cluster applications may occasionally require most of a system's processing power. Consider the demands from the application on a single member, given the combination of application resource requirements. The combination of applications is particularly important on small systems with little memory, and on large systems with many simultaneous users.

The *Introduction to VAXcluster Application Design* is intended for system designers and application programmers who are designing and coding new applications that will run on a VAXcluster system or who are migrating VMS applications to a VAXcluster system.

Real-Time Processing

Not all real-time applications are suitable for use in a cluster. The reason is that the active nodes pause whenever a system joins or leaves the cluster. These cluster-state transitions typically last only a few seconds and they occur only occasionally, but they do affect the entire cluster.

If you expect a particular system to leave the cluster frequently, perhaps because of software testing or nonstandard hardware, be aware that each absence will cause a brief processing pause on the other nodes in the VAXcluster system. Designing a real-time application that can tolerate these pauses requires a knowledgeable VMS system programmer.

A VAXcluster environment can be used effectively with real-time applications when a fault tolerant system is dedicated to real-time transaction processing or data collection. This system could use a network connection to carry output from its applications to a VAXcluster system for further, less time-critical processing. In this way, real-time applications can be shielded from any fluctuation in processing that the VAXcluster system experiences.

Fault Tolerance

The ability of a system to continue automatically processing applications even if a hardware component fails is called **fault tolerance**. Although the VAXcluster hardware is designed for high availability, the VAXcluster software does not implement a fault-tolerant system. If a processor fails, the processes on that processor do not automatically continue running on another processor.

It is possible to design a fault-tolerant application to run on a VAXcluster system. A CI only cluster is the best system for such an application, because the CI hardware is designed for high availability. Another course, **VAXcluster System and Application Design**, can help you decide whether your application can be made fault-tolerant.

Applications and Cluster Design

Clusters should be designed around their applications so that they perform best the things that they will be doing most of the time. Configuration performance and efficiency is always dependent on what the system is actually doing. If the demands of the application (expected cluster workload) are not taken into account before a system is designed and configured, then there is no guarantee that the system will perform adequately. Therefore, knowing cluster features and your primary application features is the key to a successful cluster configuration.

Combining Applications

To plan the appropriate combination of applications for each cluster member:

- Determine what software will be used on which nodes, grouping compatible images where possible.
- Adjust planned capacity to include expected users, including distribution of user population.
- Consider these issues with system security and ease-of-management issues.
- Decide which members will be part of the cluster and which will be connected by DECnet software.

Building a VAXcluster System

INTRODUCTION

Building a VAXcluster system involves preparing the cluster environment and then satisfying the cluster specification derived in the preparation phase by executing the activities necessary to build the cluster.

This means installing, initializing, and starting the system components. A VAXcluster system is created when the first system boots. This module outlines the procedures for building clusters of all types, and then looks more closely at the operations, in order, that must be conducted to build a cluster.

OBJECTIVES

After completing this module, students should be able to:

- Initialize hardware
- Install VMS software
- Install all system and product license keys
- Configure the DECnet network and assign a cluster alias when appropriate
- Add members to the cluster
- Set all members to automatically start up in the cluster
- Configure system disks
- Set up VMSMAIL_PROFILE.DATA and User Authorization Files
- Set up print and batch queues
- Manage disk and tape volumes
- Set up a LAT environment

RESOURCES

- *VMS VAXcluster Manual*
- *VMS Networking Manual*
- *VMS System Dump Analyzer Utility Manual*
- *Guide to Setting Up a VMS System*
- *Guide to Maintaining a VMS System*
- *VMS Access Control List Editor Manual*
- *HSC User Guide*

BUILDING A VAXcluster SYSTEM

Before you configure a VAXcluster system, you should be aware of the following hardware considerations:

- Hardware must be brought up to the required, compatible revision levels
- Each CI node must have a unique CI port number

Steps Needed to Build a New VAXcluster System

Once the hardware is properly set up, you can begin to build the VAXcluster software environment.

- Initialize HSC and ISE units
  - Customer Service should set up the hardware
  - HSC node names and ids may need to be modified
  - ISE node names, ids, and unit numbers may need to be modified
- Set up boot command procedures, if you can do so without VMS running.
- Install the latest version of the VMS operating system on one node, the first VMS node in this cluster. In a local area cluster, the installation is done on a boot server. In a CI cluster, the installation is done on any VMS node connected to the CI bus.
- Install the software licenses for the VMS operating system, VAXcluster software, DECnet software, and any other software you are licensed to use.
- Install layered products
- Configure the DECnet database @sys$manager:NETCONFIG.COM
- Invoke CLUSTER_CONFIG.COM uit sys$startup:
  - Add and remove VMS nodes
  - Change configurations
  - Create system disks, quorum disks
- Set up authorization files, Mail file, and startup command procedures sysuaf.dat, netproxy.dat rightslist.dat vms_mail_profile.data jbcsysque.dat alle opstartprocedures
- Set up boot command procedures, if you did not before
- Configure system disks
- Set up the operating environment

aanpassen sys$specific:[sysmgr]modparams.dat o.a. expected votes

INITIALIZING THE HSC AND ISE UNITS

HSC Installation Procedure

This example shows the procedure for installation in a cluster with full failover capabilities. Installation in a cluster without failover capabilities requires a shutdown of the cluster. After shutdown, continue with step 4, shown below. For full details, see the *HSC User Guide*.

Example 5–1 Installing the HSC Unit in a Cluster with Failover Capabilities

```
HSC> SHO SYS
17-April-1989 10:32:14.02 Boot: 17-March-1989 11:14:34.14 Up: 3:11
Version: V370  System ID: %X003 Name: CLOWN
Front Panel: Enabled Sector Size: 512
Console Dump: Enabled Load Dump: Disabled
Restart: Warm
Automatic DITs: Enabled  Interval = 1
Disk Allocation Class: 1 Tape Allocation Class: 5
Startup Command File: Disabled
Maximum Tape Drives: 5
Maximum Formatters: 1
SETSHO-I Program Exit
```

1. Enter the SHO SYS command at the HSC> prompt.
2. After printing a hard copy of the above information, failover disk drives to a single HSC, making sure that no disk or tape drives are on-line to this HSC unit.
3. After failover, press the ONLINE switch to the out position.
4. Open the HSC front panel.
   - Remove all old media
   - Insert new media
   - Write-enable new media
5. Press the INIT and FAULT buttons simultaneously.
6. Release the INIT button.
7. Continue pressing the FAULT button until you receive the following console message:

   ```
   INIPIO-I Booting
   ```
8. Release the FAULT button.

9. After the HSC is booted, enter CTRL/Y, then enter the RUN SETSHO command at the HSC> prompt.

10. Set the following parameters at the SETSHO> prompt, referring to the printed copy from the SHO SYS command executed earlier:

```
SETSHO>SET NAME CLOWN
SETSHO>SET ID %X003
SETSHO>SET ALLOCATE DISK 1
SETSHO>SET ALLOCATE TAPE 5
SETSHO>EXIT
SETSHO-Q Rebooting HSC; Y to continue, CTRL/Y to abort:? Y

INIPIO Booting...

HSC Version V370  5-Apr-1989 22:28:51    System CLOWN
```

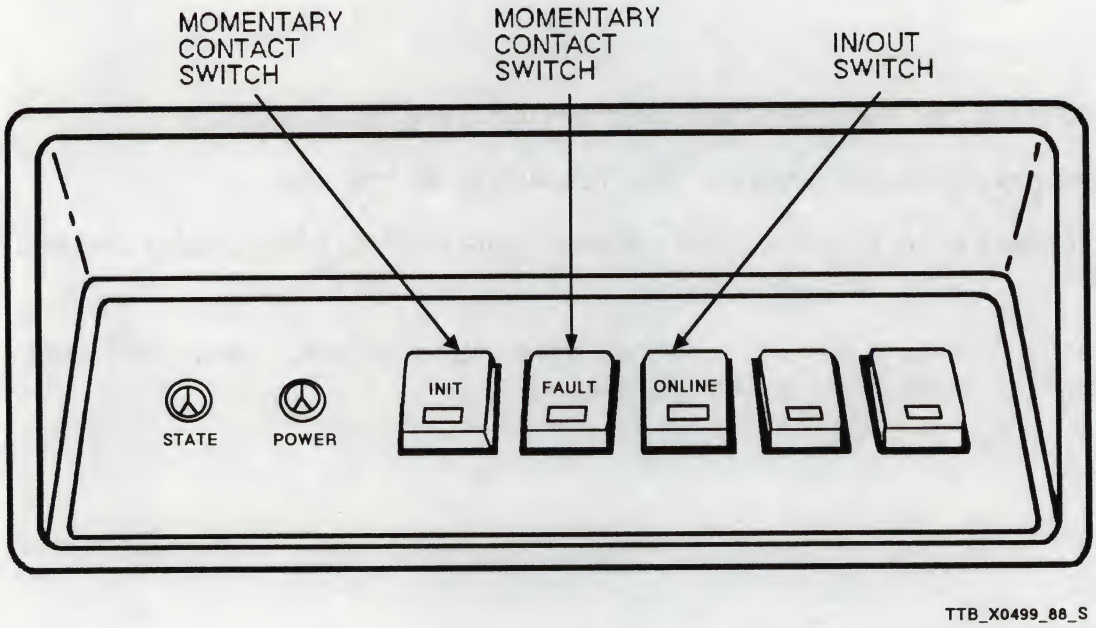
11. Issue another SHO SYS command and compare the old and new parameters.

12. If all parameters are correct, press the ONLINE switch to the in position.

13. Failover the disks to this HSC and install software on the alternate HSC following the same procedure.

If you return an HSC to the cluster with a different NAME, ID, or allocation class than it used previously, you must reboot all of the VAX systems.

Figure 5–1 HSC Front Panel



Modifying HSC System Parameters

From the HSC console terminal:

- Type CTRL/Y to get the HSC> prompt
- Run the SETSHO utility
- Use SETSHO commands to change parameter defaults:
  - Set SCS node name

    ```
    SETSHO>SET NAME nnnnnn
    ```

  - Set SCS ID number

    ```
    SETSHO>SET ID %Xnnnnnn
    ```

  - Set disk allocation class

    ```
    SETSHO>SET ALLOCATE DISK nnn
    ```

  - Set tape allocation class

    ```
    SETSHO>SET ALLOCATE TAPE nnn
    ```

  - Enter the SHO SYS command to display system parameter values
  - Enter the EXIT command to leave SETSHO

The HSC unit can be accessed from a VMS node running FYDRIVER.

```
$RUN SYS$SYSTEM:SYSGEN
SYSGEN> CONNECT FYA0/NOADAPTER   !load fydriver (HSC driver)
SYSGEN> EXIT
$SET HOST/HSC hscname
HSC>RUN SETSHO
SETSHO> SET NAME nnnnn           !as above
SETSHO> EXIT                     !exit from SETSHO
SETSHO-Q Rebooting HSC; Y to continue,  CTRL/Y to abort:? Y

INIPIO Booting...

HSC Version 370  5-Apr-1989 22:28:51    System CLOWN
```

REMEMBER

After changing any of the above parameters, the HSC unit will automatically reboot.

INITIALIZING ISE UNITS

Setting ISE Parameters

Set host directly to the ISE to:

- Set parameters
- Execute diagnostics
- Execute utilities

To set host to the ISE and get a directory listing:

```
$ MCR SYSGEN CONNECT FYA0/NOADAPTER                    ! load DUP driver
$ SET HOST/DUP/SERVER=MSCP$DUP/TASK=DIRECT R3QRNQ
%HSCPAD-I-LOCPROGEXE, Local program executing - type ^\ to exit
Copyright 1988 Digital Equipment Corporation
DRVEXR V1.1 D 1-MAR-1989 12:11:13
DRVTST V1.1 D 1-MAR-1989 12:11:13
HISTRY V1.0 D 1-MAR-1989 12:11:13
ERASE  V1.3 D 1-MAR-1989 12:11:13
PARAMS V1.2 D 1-MAR-1989 12:11:13
DIRECT V1.0 D 1-MAR-1989 12:11:13
End of directory

%HSCPAD-S-REMPGMEND, Remote program terminated - message code 1.
%HSCPAD-S-END, Control returned to node LARRY
```

Invoking PARAMS

Local programs are invoked by using the console commands for the MicroVAX 3000 series or through the VMS operating system using the command:

```
$ SET HOST/DUP/SERVER=MSCP$DUP/TASK=progname nodename
```

PARAMS is invoked in the same manner and, once executing, all interaction is through the use of commands and responses.

Table 5–1 Valid PARAMS Commands and Functions

Command	Definition
HELP	Shows all PARAMS commands and their syntax
SET	Sets a parameter to a value
SHOW	Displays a parameter or a class of parameters
STATUS	Displays module configuration, history, or current counters
WRITE	Records the device parameters you changed using the SET command
EXIT	Terminates the PARAMS local program

VMS INSTALLATION

VMS installation is fundamentally the same for standalone and clustered systems. The specific cluster concerns involve allocation classes and security matters.

After Installing VMS Software

- Determine disk and boot server allocation classes
- Specify passwords
- Configure devices
- Invoke AUTOGEN
- Reboot the system
- Install layered products

Standalone backup system disk!

INSTALLING REQUIRED SOFTWARE LICENSE KEYS

The following types of software have license keys:

- VMS operating system
  - Separate license for each VMS member
- System Integrated Products (SIPS)
  - VAXcluster software
  - DECnet VAX software

    Endnode (DVNETEND)

    Router (DVNETRTG)
- Other layered products

License Management Facility (LMF) and PAKs

The LMF allows cluster-wide management of licenses through Product Authorization Keys (PAKs) issued by Digital for specific products and system configurations.

The VMS license PAK is a special case:

Niet te delen met een ander systeem

- Has a NO_SHARE attribute
  - It cannot be shared between members
  - Every VMS system must be licensed for a specific processor type
  - The PAK must be specific for each member in the cluster
- Interactive users cannot log in, except at the console, without a VMS license PAK enabled on that member

Installation of the VMS License Key

If you are **NOT** a VMS Service Customer:

- Use the VMSLICENSE procedure @sys$update:vmslicense

  (or)

- Use the LICENSE REGISTER command

If you **ARE** a VMS Service Customer:

- Install VMS software with a W-KIT
- Your VMS Service Update PAK (SUP) is generated when you apply the mandatory update (MUP) after installation
- The command procedure SYS$UPDATE:LMF$CONFIG.COM creates a small, system-specific LICENSE database
  - SYS$SPECIFIC:[SYSEXE]LMF$SYSTEM.LDB
  - Determines appropriate VMS SUP for your system
  - Registers it in this database

CONFIGURING THE DECnet VAX NETWORK

This process typically includes the following operations:

- Invoking the NETCONFIG.COM command procedure
- Making remote node data available cluster-wide
- Optionally defining an alias node identifier for the cluster
- Starting the network

After installing VMS software and required licenses:

- Invoke the NETCONFIG.COM command procedure, entering information about your node when prompted, and responding YES when the procedure asks whether you want to configure the network
- Provide information about your system, as prompted
- Establish default DECnet accounts
  - Mail
  - FAL
  - PHONE
  - NML
  - Default account with no access for task 0 objects
- If you are going to define a cluster alias, respond NO when asked if you wish to start the network

Example 5–2 shows typical responses for a cluster network configuration session using NETCONFIG.COM.

Example 5–2 Sample Interactive Network Configuration Session (Sheet 1 of 3)

```
Username: SYSTEM
Password:
 Welcome to VAX/VMS version V5.3 on node BAILEY
    Last interactive login on Monday, 26-APR-1990 09:18
    Last non-interactive login on Monday, 26-APR-1990 09:06

$ @NETCONFIG.COM

 DECnet VAX network configuration procedure

This procedure will help you define the parameters needed to get DECnet
running on this machine.  You will be shown the changes before they are
executed, in case you wish to perform them manually.

What do you want your DECnet node name to be?          [BAILEY]:          = SCS NODE
What do you want your DECnet address to be?               [1.3]:
Do you want to operate as a router?            [NO (nonrouting)]:
Do you want a default DECnet account?                      [NO]: RETURN   niet meer nodig
Do you want to use the CI as a DECnet datalink?            [NO]: RETURN
Do you want a default account for the MAIL object?        [YES]: RETURN
Do you want a default account for the FAL object?          [NO]: RETURN   FAL $ SERVER
Do you want a default account for the PHONE object?       [YES]: RETURN
Do you want a default account for the NML object?          [NO]: RETURN
```

Example 5–2 Sample Interactive Network Configuration Session (Sheet 2 of 3)

```
 Here are the commands necessary to setup your system.

$ RUN SYS$SYSTEM:NCP
    PURGE EXECUTOR ALL
    PURGE KNOWN LINES ALL
    PURGE KNOWN CIRCUITS ALL
    PURGE KNOWN LOGGING ALL
    PURGE KNOWN OBJECTS ALL
    PURGE MODULE CONFIGURATOR KNOWN CIRCUITS ALL
$ DEFINE/USER SYS$OUTPUT NL:
$ DEFINE/USER SYS$ERROR NL:
$ RUN SYS$SYSTEM:NCP ! Remove existing entry, if any
    PURGE NODE 1.3 ALL
    PURGE NODE BAILEY ALL
$ RUN SYS$SYSTEM:NCP
    DEFINE EXECUTOR ADDRESS 1.3 STATE ON
    DEFINE EXECUTOR NAME BAILEY
    DEFINE EXECUTOR MAXIMUM ADDRESS 1023
    DEFINE EXECUTOR TYPE NONROUTING IV
    DEFINE OBJECT TASK NUMBER 0 USER ILLEGAL PASSWORD DISABLED
    DEFINE OBJECT MAIL NUMBER 27 USER MAIL$SERVER PASSWORD annzathwas
$ DEFINE/USER_MODE SYSUAF SYS$SYSTEM:SYSUAF.DAT
$ RUN SYS$SYSTEM:AUTHORIZE
    ADD MAIL$SERVER /OWNER="MAIL$SERVER DEFAULT" -
 /PASSWORD=annzathwas -
 /UIC=[376,374] /ACCOUNT=DECNET -
 /DEVICE=SYS$SPECIFIC: /DIRECTORY=[MAIL$SERVER] -
 /PRIVILEGE=(TMPMBX,NETMBX) -
 /DEFPRIVILEGE=(TMPMBX,NETMBX) -
 /FLAGS=(NOCAPTIVE,RESTRICTED,NODISUSER) /LGICMD=NL: -
 /NOBATCH /NOINTERACTIVE
    MODIFY MAIL$SERVER /PASSWORD=annzathwas
$ CREATE/DIRECTORY SYS$SPECIFIC:[MAIL$SERVER] /OWNER=[376,374]
$ RUN SYS$SYSTEM:NCP
    DEFINE OBJECT PHONE NUMBER 29 USER PHONE$SERVER PASSWORD kaybpyur
$ DEFINE/USER_MODE SYSUAF SYS$SYSTEM:SYSUAF.DAT
$ RUN SYS$SYSTEM:AUTHORIZE
    ADD PHONE$SERVER /OWNER="PHONE$SERVER DEFAULT" -
 /PASSWORD=kaybpyur -
 /UIC=[376,372] /ACCOUNT=DECNET -
 /DEVICE=SYS$SPECIFIC: /DIRECTORY=[PHONE$SERVER] -
 /PRIVILEGE=(TMPMBX,NETMBX) -
 /DEFPRIVILEGE=(TMPMBX,NETMBX) -
 /FLAGS=(NOCAPTIVE,RESTRICTED,NODISUSER) /LGICMD=NL: -
 /NOBATCH /NOINTERACTIVE
    MODIFY PHONE$SERVER /PASSWORD=kaybpyur
$ CREATE/DIRECTORY SYS$SPECIFIC:[PHONE$SERVER] /OWNER=[376,372]
```

Example 5-2 Sample Interactive Network Configuration Session (Sheet 3 of 3)

```
$ RUN SYS$SYSTEM:NCP
    DEFINE LINE    BNA-0 STATE ON
    DEFINE CIRCUIT BNA-0 STATE ON COST 4
    DEFINE LOGGING MONITOR STATE ON
    DEFINE LOGGING MONITOR EVENTS 0.0-9
    DEFINE LOGGING MONITOR EVENTS 2.0-1
    DEFINE LOGGING MONITOR EVENTS 4.2-13,15-16,18-19
    DEFINE LOGGING MONITOR EVENTS 5.0-18
    DEFINE LOGGING MONITOR EVENTS 128.0-4

Do you want these commands to be executed?  [YES]:

%NCP-I-SUCCESS,  Success
    .
    .
    .
The changes have been made.
```

If you have not already registered the DECnet VAX license key, do so now.

- After the key has been registered, invoke the command procedure SYS$MANAGER:STARTNET.COM to start up DECnet VAX with these changes.
- If the key is already registered, and you wish to define an alias node identifier for the cluster, do not start the network yet:

```
Do you want DECnet started? [YES]: NO
```

Defining Cluster Alias Operations

To define a cluster alias:

- Execute the Network Control Program (NCP)
- Define the cluster alias in the common database

For example:

```
$ RUN SYS$SYSTEM:NCP
NCP> DEFINE NODE 1.10 NAME CIRCUS
NCP> DEFINE EXECUTOR ALIAS NODE CIRCUS
     ! If the node has not started the network, include this line:  SET EXEC STATE OFF
NCP> DEFINE EXECUTOR ALIAS INCOMING ENABLED
NCP> EXIT                                                @ STARTNET
$
```

Information specified here:

- Is entered in the DECnet VAX permanent executor database
- Takes effect when you start the network

Enabling a Cluster Alias

If you have defined an alias node identifier for your cluster, you can enable alias operations for other cluster nodes after the nodes have joined the cluster.

To enable a cluster alias:

On each node:

- Log in as system manager and execute the NCP utility.

```
$SET PROCESS/PRIVILEGES=(OPER,SYSPRV)
$RUN SYS$SYSTEM:NCP
NCP> SET EXECUTOR STATE OFF
NCP> DEFINE EXECUTOR ALIAS INCOMING ENABLED
NCP> EXIT
$ @SYS$MANAGER:STARTNET.COM
```

Starting the Network

If you chose to define a cluster alias, you answered NO to the NETCONFIG question to start the network. You will need to start the network manually.

To start the network:

- Invoke STARTNET.COM

```
$ @SYS$MANAGER:STARTNET.COM
```

- To ensure that the network is started each time the system boots, add the following to your site-specific startup command procedure:

```
$ @SYS$MANAGER:STARTNET.COM
```

Copying Remote Node Databases

Some sites with large networks maintain remote node data in a central database file.

If you want to make the data available cluster-wide:

- Make sure the network is up
- Run NCP
- Copy remote node database entries from that central file
- For example, if the file resides on CURLY, copy entries from the volatile database on CURLY to the permanent database on your system disk. Then update the volatile database:

```
$ RUN SYS$SYSTEM:NCP
NCP> SET NODE 1.93 NAME CURLY       ! assume we know this somehow
NCP> COPY KNOWN NODES FROM CURLY USING VOLATILE TO PERMANENT
NCP> SET KNOWN NODES ALL
```

- Only node names and addresses are copied

Using the CI Bus for DECnet Circuits

Where Ethernet is not available, the CI bus can be used for DECnet traffic.

To use the CI bus:

- Place the following commands in the SYS$MANAGER:SYCONFIG.COM on all nodes. (In this example, the nodes are BARNUM, RNGLNG, and BAILEY.)

```
$ RUN SYS$SYSTEM:SYSGEN
SYSGEN> CONNECT CNA0/NOADAPTER
SYSGEN> EXIT
```

- Execute the following commands on a routing node (in this example, BARNUM) to change its permanent database so the router will recognize the other nodes in the VAXcluster system:

```
$ RUN SYS$SYSTEM:NCP
NCP> DEFINE LINE CI-0 STATE ON
NCP> DEFINE CIRCUIT CI-0.1 TRIBUTARY 1 STATE ON
NCP> DEFINE NODE 1.2 NAME RNGLNG
NCP> DEFINE CIRCUIT CI-0.2 TRIBUTARY 2 STATE ON
NCP> DEFINE NODE 1.3 NAME BAILEY
```

NOTE

Each number 1 in the DEFINE CIRCUIT command corresponds to RNGLNG's CI port number. Each number 2 in the DEFINE CIRCUIT command corresponds to BAILEY's CI port number. Use additional DEFINE CIRCUIT and DEFINE NODE commands to add other nodes to the network.

- Execute these commands on a non-routing node (in this example, BAILEY) to change its permanent database so it will recognize the routing node(s) in its VAXcluster system:

```
RUN SYS$SYSTEM:NCP
NCP> DEFINE LINE CI-0 STATE ON
NCP> DEFINE CIRCUIT CI-0.0 TRIBUTARY 0 STATE ON
NCP> DEFINE NODE 1.1 NAME BARNUM
```

NOTE

Each number 0 in the DEFINE CIRCUIT command corresponds to BARNUM's CI port number. Use additional DEFINE NODE commands to add other nodes to the network.

- Start up DECnet VAX software by typing:

```
@SYS$MANAGER:STARTNET.COM
```

USING CLUSTER_CONFIG.COM

- CLUSTER_CONFIG is a command procedure that automates most of the common operations needed to build a VAXcluster system.
- HELP is available at all prompts by entering "?"

CLUSTER_CONFIG.COM Functions

- Change a standalone system into a cluster system
- Add a node to the cluster
- Remove a node from the cluster
- Change a cluster node's characteristics
- Create a system disk with the same common files, but without any system roots

CAUTION

You may not initiate concurrent CLUSTER_CONFIG sessions in the same cluster.

Table 5–2 Summary of CLUSTER_CONFIG.COM Functions

Function	Operations Performed
ADD	Establishes the new node's root directory on a cluster common system disk
	Generates the node's system parameter files (VAXVMSSYS.PAR and MODPARAMS.DAT) in its SYS$SPECIFIC:[SYSEXE] directory
	Updates the permanent and volatile remote node network databases for the system
	If the new node is a satellite, updates SYS$MANAGER:NETNODE_UPDATE.COM on the local system
	Generates the new node's page and swap files (PAGEFILE.SYS and SWAPFILE.SYS)
	Optionally sets up a cluster quorum disk
	Sets the allocation class (ALLOCLASS) value for the new node, if the node is being added as a disk server
	Generates an initial (temporary) startup procedure for the new node that: • Invokes NETCONFIG.COM to configure the network • Invokes AUTOGEN to set appropriate SYSGEN parameter values for the node • Invokes LMF$CONFIG.COM to create the VMS license • Shuts down and reboots the node with normal startup procedures
REMOVE	Deletes another node's root directory from the local system's system disk
	If the node is a satellite, updates SYS$MANAGER:NETNODE_UPDATE.COM on the local system
	Updates the permanent and volatile remote node network databases on the local system
CHANGE	Enables or disables the local system as a disk server
	Enables or disables the local system as a boot server
	Enables or disables the Ethernet for cluster communications on the local system
	Enables or disables a quorum disk on the local system
	Changes the local system's ALLOCLASS value
	Changes a satellite's Ethernet hardware address
CREATE	Duplicates the local system's system disk
	Removes all "added" system roots from the new disk.

Building a Local Area VAXcluster Environment

The steps needed to build a local area VAXcluster system are:

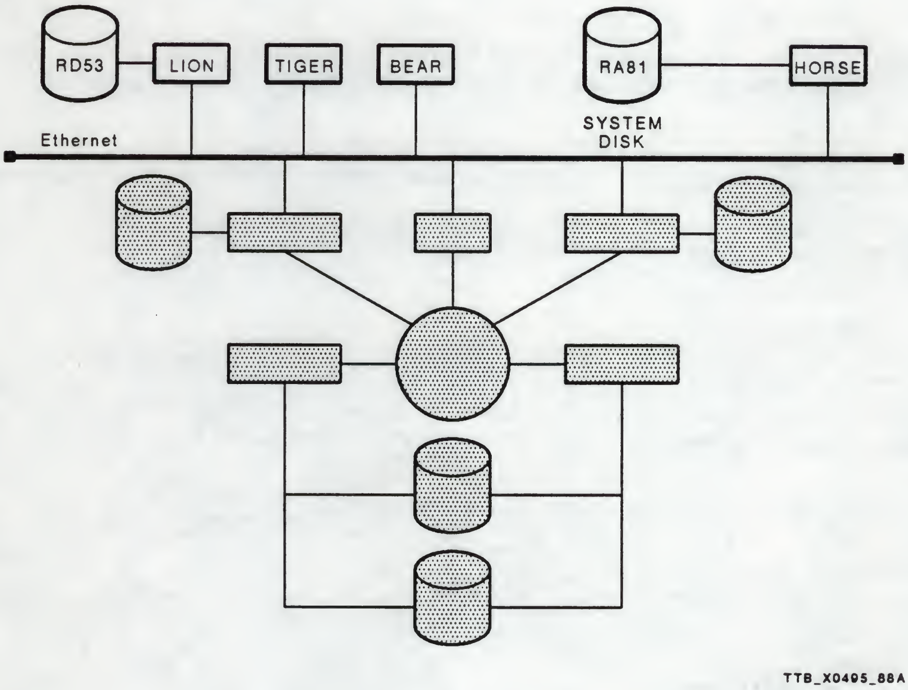
- Install or upgrade VMS software on the system disk of the boot node
- Configure the boot node during installation, upgrade, or CLUSTER_CONFIG.COM
- Configure DECnet software and start the network
- Create system roots for satellites with the ADD function in CLUSTER_CONFIG

Building a Local Area VAXcluster System with One Boot Server

Figure 5–2 assumes the following:

- There is a local area cluster with one MicroVAX II boot server (HORSE), two VAXstation 2000 satellites (TIGER, BEAR), and one VAXstation II/GPX workstation (LION).
- TIGER and BEAR have no local system disk.
- LION has an RD53 disk drive.

Figure 5–2 A Local Area VAXcluster System with a Single Boot Server



Using the ADD function

Example 5-3 Sample Interactive CLUSTER_CONFIG Session to Add a Satellite Node with Local Page and Swap Files (Sheet 1 of 3)

```
$ @CLUSTER_CONFIG
                    Cluster Configuration Procedure

    Use CLUSTER_CONFIG to set up or change a VAXcluster configuration.
    To ensure that you have the required privileges, invoke this procedure
    from the system manager's account.

    Enter ? for help at any prompt.

            1. ADD a node to the cluster.
            2. REMOVE a node from the cluster.
            3. CHANGE a cluster node's characteristics.
            4. CREATE a second system disk for HORSE.

    Enter choice [1]: [RETURN]

    The ADD function adds a new node to the cluster.

    If the node being added is a voting member, EXPECTED_VOTES in all
    other cluster members' MODPARAMS.DAT must be adjusted, and the
    cluster must be rebooted.

    If the new node is a satellite, the network databases on HORSE are
    updated. The network databases on all other cluster members must be
    updated.

    For instructions, see the VMS VAXcluster Manual.

What is the node's DECnet node name? LION
What is the node's DECnet address? 1.99
Will LION be a satellite [Y]? [RETURN]
Verifying circuits in network database...

    This procedure will now ask you for the device name of LION's system root.
    The default device name (DISK$VAXVMSRL5:) is the logical volume name of
    SYS$SYSDEVICE:.

What is the device name for LION'S system root [DISK$VAXVMSRL5:]? [RETURN]
What is the name of the new system root [SYS10]? [RETURN]    Satelliet begint altijd by 10
Allow conversational bootstraps on LION [NO]? [RETURN]
The following workstation windowing options are available:

      1. No workstation software
      2. VWS Workstation Software
      3. DECwindows Workstation Software

Enter choice [1]: 2

Creating directory tree SYS10...
%CREATE-I-CREATED, $1$DUA0:<SYS10> created
%CREATE-I-CREATED, $1$DUA0:<SYS10.SYSEXE> created
   .
   .
   .
System root SYS10 created.
```

Example 5-3 Sample Interactive CLUSTER_CONFIG Session to Add a Satellite Node with Local Page and Swap Files (Sheet 2 of 3)

```
Will LION be a disk server [N]? [RETURN]      MSCP_LOAD wordt al of niet gezet.
What is LION's Ethernet hardware address? 08-00-33-41-77-9F
Updating network database...
Size of pagefile for LION [10000 blocks]? 20000
Size of swap file for LION [8000 blocks]? 12000
Will a local disk on LION be used for paging and swapping? YES
Creating temporary page file in order to boot LION for the first time...
%SYSGEN-I-CREATED, $1$DUA0:<SYS10.SYSEXE>PAGEFILE.SYS;1 created

    This procedure will now wait until LION joins the cluster.

    Once LION joins the cluster, this procedure will ask you
    to specify a local disk on LION for paging and swapping.

    Please boot LION now.

Waiting for LION to boot...
    .
    .
    .
```

At this point, the user should enter the appropriate boot command at the satellite's console prompt [>>>]:

- For MicroVAX II, VAXstation II, and MicroVAX 3000 series satellites:
  >>> B XQ
- For MicroVAX 2000 and VAXstation 2000 satellites:
  >>> B ES

Example 5–3 Sample Interactive CLUSTER_CONFIG Session to Add a Satellite Node with Local Page and Swap Files (Sheet 3 of 3)

```
    .
    .
    .
The local disks on LION are:

Device            Device       Error     Volume     Free     Trans   Mnt
Name              Status       Count     Label      Blocks   Count   Cnt
LION$DUA0:      Online           0

Which disk can be used for paging and swapping? LION$DUA0:
May this procedure INITIALIZE LION$DUA0: [YES]? NO
Mounting LION$DUA0:...
PAGEFILE.SYS already exists on LION$DUA0:

***************************************

Directory LION$DUA0:[SYS0.SYSEXE]

PAGEFILE.SYS;1      23600/23600

Total of 1 file, 23600/23600 blocks.

***************************************

What is the file specification for the page file on
LION$DUA0: [ <SYS0.SYSEXE>PAGEFILE.SYS ]? RETURN
%CREATE-I-EXISTS, LION$DUA0:<SYS0.SYSEXE> already exists
This procedure will use the existing pagefile,
LION$DUA0:<SYS0.SYSEXE>PAGEFILE.SYS;.

SWAPFILE.SYS already exists on LION$DUA0:

***************************************

Directory LION$DUA0:[SYS0.SYSEXE]

SWAPFILE.SYS;1      12000/12000

Total of 1 file, 12000/12000 blocks.

***************************************

What is the file specification for the swap file on
LION$DUA0: [ <SYS0.SYSEXE>SWAPFILE.SYS ]? RETURN
This procedure will use the existing swapfile,
LION$DUA0:<SYS0.SYSEXE>SWAPFILE.SYS;.

     AUTOGEN will now reconfigure and reboot LION automatically.
     These operations will complete in a few minutes, and a
     completion message will be displayed at your terminal.

The configuration procedure has completed successfully.
```

Using Multiple Boot Servers in a Local Area VAXcluster Configuration

Figure 5–3 shows a local area VAXcluster system with two boot servers, BEAR and HORSE.

To configure this cluster, follow these steps:

1. Install the latest version of the VMS operating system on a boot server, for example, HORSE.
   - This simultaneously upgrades all satellites that boot from the server.
2. Generate a second system disk on BEAR in one of two ways:
   - Install VMS operating system software on BEAR.
     (All layered products would have to be installed on this system disk.)
   - Use CLUSTER_CONFIG.COM CREATE:
     — Attach the disk that will eventually be BEAR's system disk to a disk port on HORSE. (It is not connected to BEAR).

     — Mount BEAR's system disk on HORSE.

     — Use the CLUSTER_CONFIG CREATE command to copy HORSE's system disk and remove all roots.

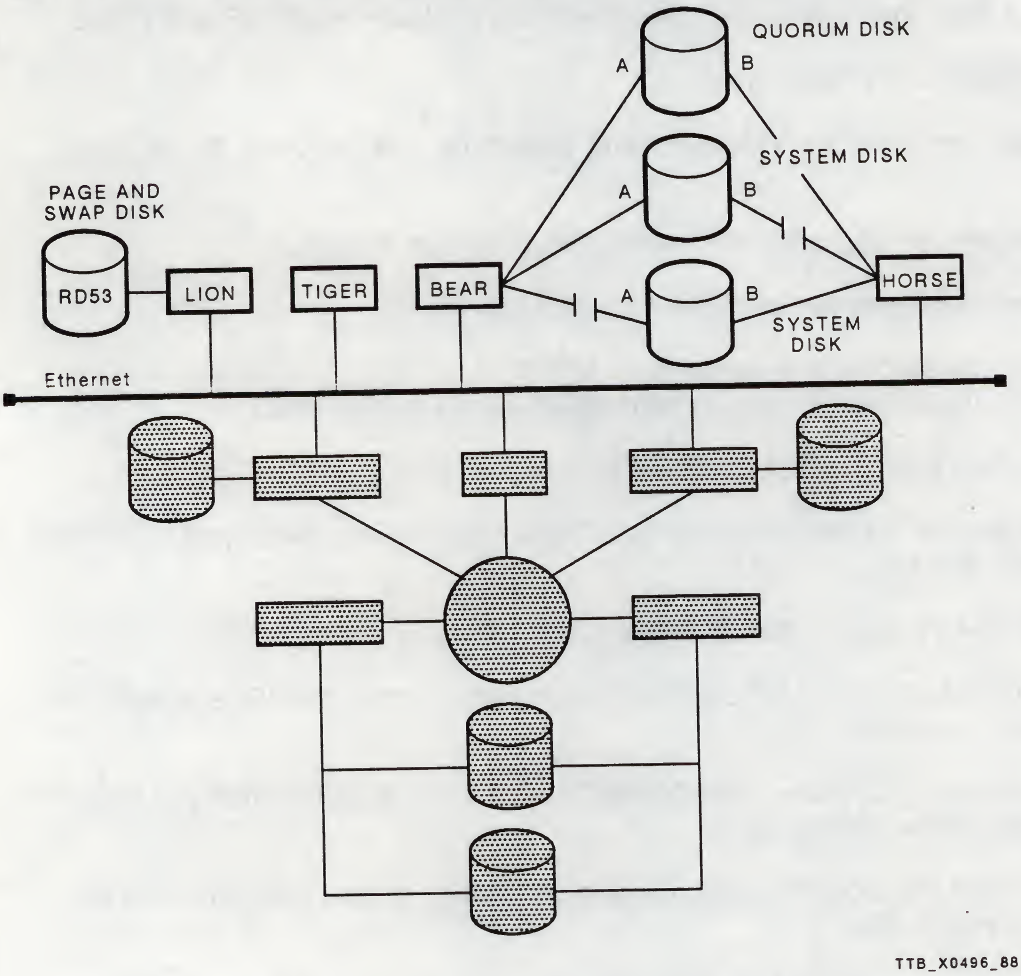
     All software in HORSE's VMS$COMMON directory tree is now available to BEAR, with none of the system roots.

     — Use CLUSTER_CONFIG.COM, running on HORSE, to add node BEAR to the second system disk.

     This is done by specifying the new volume name, and a root on the new volume.

     — Dismount the new disk from HORSE, port it to BEAR (push in the port button on the drive to allow BEAR to access it), and boot BEAR.

Figure 5–3 A Local Area VAXcluster System with Two Boot Servers



Multiple Boot Servers (Cont.)

3. Use CLUSTER_CONFIG to change the configuration to include a quorum disk.

   - The quorum disk should be dynamically dual-ported (both port buttons pushed in) to both boot servers.
   - The quorum disk will fail over automatically if the node it is ported to fails. This will retain quorum for the remaining boot server and the satellites.
   - When you assign the two boot servers the same allocation class and put all satellite system roots on the quorum disk, rather than the boot server system disks, the satellites are assured of maximum availability.
   - The port buttons for each system disk **must** be pushed, because dual-ported system disks are not supported. Failover will work, but it must be accomplished by manually pushing the port button after a system failure (static dual-porting).

4. Change EXPECTED_VOTES on each specific system root (boot servers and satellites) to include:

   - The votes of each boot server and other VMS members with votes
   - The votes of the quorum disk

5. Reboot the cluster.

6. Use CLUSTER_CONFIG on BEAR or HORSE to add additional satellites to the new system disk.

Using the REMOVE Function

Example 5–4 illustrates the use of CLUSTER_CONFIG on node BARNUM to remove satellite node LION from the cluster.

Example 5–4 Sample Interactive CLUSTER_CONFIG.COM Session to Remove a Satellite Node with Local Page and Swap Files

```
$ @CLUSTER_CONFIG
                 Cluster Configuration Procedure

    Use CLUSTER_CONFIG to set up or change a VAXcluster configuration.
    To ensure that you have the required privileges, invoke this procedure
    from the system manager's account.

    Enter ? for help at any prompt.

            1. ADD a node to the cluster.
            2. REMOVE a node from the cluster.
            3. CHANGE a cluster node's characteristics.
            4. CREATE a second system disk for BARNUM.

    Enter choice [1]: 2

    The REMOVE function disables a node as a cluster member.

        o It deletes the node's root directory tree.

        o It removes the node's network information
          from the network database.

    If the node being removed is a voting member, you must adjust
    EXPECTED_VOTES in each remaining cluster member's MODPARAMS.DAT.
    You must then reconfigure the cluster, using the procedure described
    in the VMS VAXcluster Manual.

What is the node's DECnet node name? LION
Verifying network database...
Verifying that SYS10 is LION's root...

    WARNING - LION's page and swap files will not be deleted.
              They do not reside on $1$DUA0:.

Deleting directory tree SYS10...
%DELETE-I-FILDEL, $1$DUA0:<SYS10>SYSCBI.DIR;1 deleted (1 block)
%DELETE-I-FILDEL, $1$DUA0:<SYS10>SYSERR.DIR;1 deleted (1 block)
    .
    .
    .
System root SYS10 deleted.
Updating network database...
The configuration procedure has completed successfully.
```

Enabling a Quorum Disk

A quorum disk can be enabled during installation, or by using the CLUSTER_CONFIG CHANGE option. The following SYSGEN parameters are set on nodes that are known as **quorum disk watchers.**

Quorum disk usage is controlled by three SYSGEN parameters:

- DISK_QUORUM
  - Stores the name of the active quorum disk in ASCII format
  - A string of spaces ("   ") indicates no quorum disk (the default)
- QDSKVOTES
  - The number of votes supplied by the quorum disk (default is 1)
- QDSKINTERVAL
  - The number of seconds between quorum disk polling activity (default is 10)

Quorum disk can be:

- An HSC disk
- A DSSI disk
- A MASSBUS disk
- Local to a DSA controller (UDA, KDA, KDB)
- Used in any type of cluster

The quorum disk may **NOT** be shadowed.

System Disks Local to Ethernet Members

To build a cluster with Ethernet members that have their own system disks:

- Install VMS operating system software on the Ethernet member.
  - Answer YES to question regarding cluster membership.
  - Perform all normal tasks.

  (or)

- Invoke CLUSTER_CONFIG.COM
  - Use CREATE to copy the current cluster system disk to the new member.

    CAUTION

    The target disk on the Ethernet member is initialized and must be large enough to hold the entire system disk structure.

  - Use ADD to give the system a system root.
- Install licenses.
- Answer question in the upgrade or CLUSTER_CONFIG to make the node a cluster member.
  - Enter the group number and password.
  - Allow VAXcluster communication over the Ethernet.
- Reboot the system.
- Appropriately configure the startup files and cluster common databases.

Building a CI VAXcluster System

To build a CI cluster:

- Hook up all CI hardware
- Install or upgrade VMS software on one CI member.
- Register license keys for VMS, VAXcluster, and DECnet software.
- Invoke SYS$MANAGER:CLUSTER_CONFIG.COM to ADD additional CI nodes to the system disk.
- Modify the console media so that members boot properly (for details, see this module and the processor installation guide).
- Boot each member — the first boot of each member will invoke AUTOGEN, setting parameters selected while running CLUSTER_CONFIG. This initial boot does these things automatically, using the startup file STARTUP1.COM.
- Continue configuring cluster devices, disks, tapes, etc.
- Create or modify existing member or cluster startup command procedures related to devices, logical names, and queues.

Using Multiple System Disks

There are two methods for building a CI cluster with more than one system disk.

Method 1:

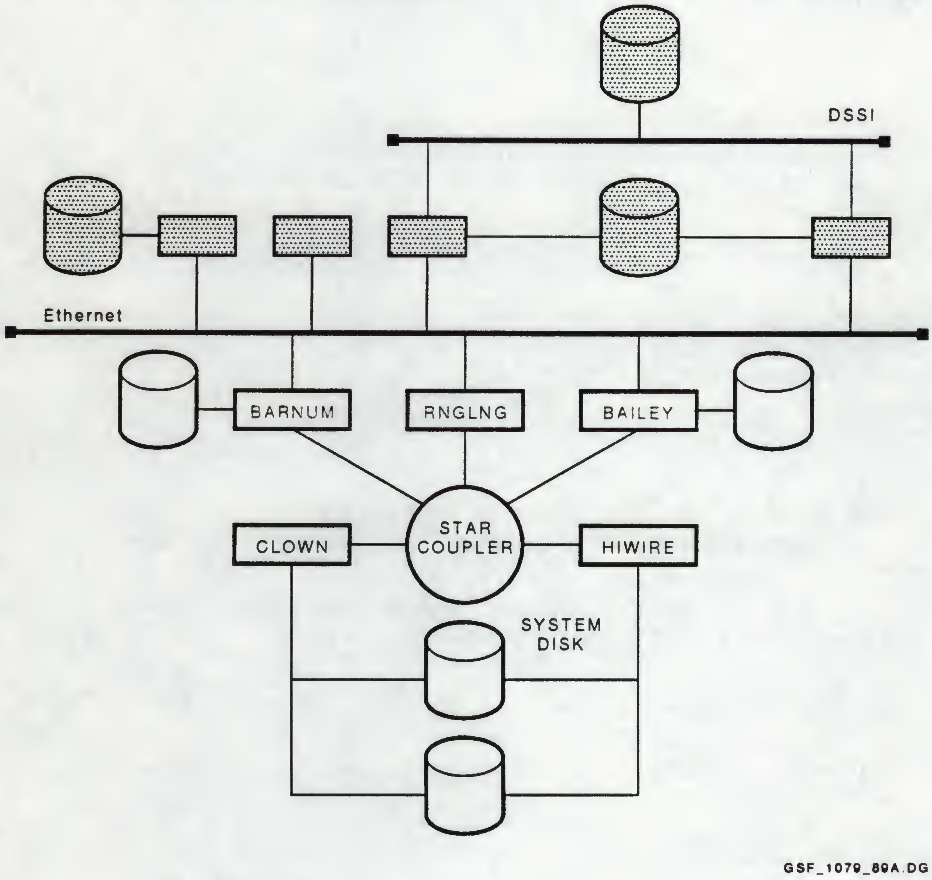
1. Install VMS software on the first member of the cluster.
2. Reboot the first member.
3. Install VMS software on a second system disk and reboot.

Method 2: Use CLUSTER_CONFIG to create a second system disk.

1. Use CREATE to copy the first system disk to an initialized disk.
   This copies the entire system environment, without any system roots.

Use ADD to add roots for CI nodes to these system disks.

Figure 5–4 A CI VAXcluster System with a Single System Disk and a Quorum Disk



Example 5-5 Sample Interactive CLUSTER_CONFIG.COM Session to Add a CI Connected Node (Sheet 1 of 2)

```
$ @CLUSTER_CONFIG
                 Cluster Configuration Procedure

    Use CLUSTER_CONFIG to set up or change a VAXcluster configuration.
    To ensure that you have the required privileges, invoke this procedure
    from the system manager's account.

    Enter ? for help at any prompt.

            1. ADD a node to the cluster.
            2. REMOVE a node from the cluster.
            3. CHANGE a cluster node's characteristics.
            4. CREATE a second system disk for BARNUM.

    Enter choice [1]: [RETURN]

    The ADD function adds a new node to the cluster.

    If the node being added is a voting member, EXPECTED_VOTES in all
    other cluster members' MODPARAMS.DAT must be adjusted, and the
    cluster must be rebooted.

    If the new node is a satellite, the network databases on BARNUM are
    updated. The network databases on all other cluster members must be
    updated.

    For instructions, see the VMS VAXcluster Manual.

What is the node's DECnet node name? BAILEY
What is the node's DECnet address? 1.3
Will BAILEY be a satellite [Y]? N
Will BAILEY be a boot server [Y]? [RETURN]

    This procedure will now ask you for the device name of BAILEY's system root.
    The default device name (DISK$VAXVMSRL5:) is the logical volume name of
    SYS$SYSDEVICE:.

What is the device name for BAILEY's system root [DISK$VAXVMSRL5:]? [RETURN]
What is the name of the new system root [SYS2]? [RETURN]
Creating directory tree SYS2...
%CREATE-I-CREATED, $1$DUA0:<SYS2> created
%CREATE-I-CREATED, $1$DUA0:<SYS2.SYSEXE> created
    .
    .
    .
System root SYS2 created.
Enter a value for BAILEY's ALLOCLASS parameter: 1
Does this cluster contain a quorum disk [N]? [return]
Updating network database...
Size of page file for BAILEY [10000 blocks]? 50000
Size of swap file for BAILEY [8000 blocks]? 20000
Will a local (non-HSC) disk on BAILEY be used for paging and swapping? N

      If you specify a device other than DISK$VAXVMSRL5: for BAILEY's
      page and swap files, this procedure will create PAGEFILE_BAILEY.SYS
      and SWAPFILE_BAILEY.SYS in the <SYSEXE> directory on the device
      you specify.
```

Example 5-5 Sample Interactive CLUSTER_CONFIG.COM Session to Add a CI Connected Node (Sheet 2 of 2)

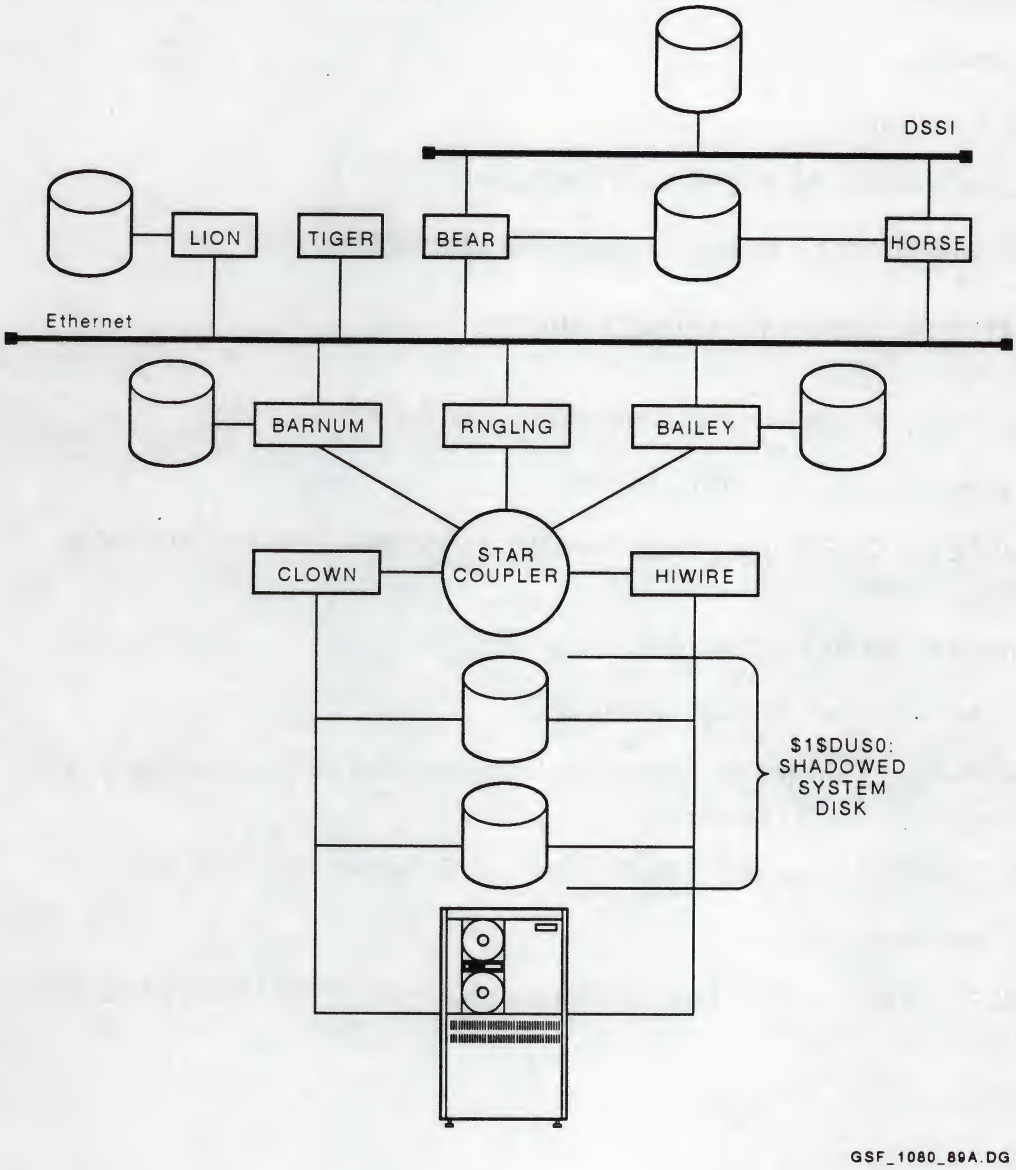
```
What is the device name for the page and swap files
[DISK$VAXVMSRL5:]? [RETURN]
%SYSGEN-I-CREATED, $1$DUA0:<SYS2.SYSEXE>PAGEFILE.SYS;1 created
%SYSGEN-I-CREATED, $1$DUA0:<SYS2.SYSEXE>SWAPFILE.SYS;1 created
The configuration procedure has completed successfully.

     BAILEY has been configured to join the cluster.

     Before booting BAILEY, you must create a new default bootstrap
     command procedure for BAILEY. See your processor-specific
     installation and operations guide for instructions.

     The first time BAILEY boots, NETCONFIG.COM and
     AUTOGEN.COM will execute automatically.

     The following parameters have been set for BAILEY:

               VOTES = 1
               QDSKVOTES = 1

     After BAILEY has booted into the cluster, you must increment
     the value for EXPECTED_VOTES in every cluster member's
     MODPARAMS.DAT. You must then reconfigure the cluster, using the
     procedure described in the VMS VAXcluster Manual.
```

Building a Mixed-Interconnect VAXcluster System

To build an MI (Mixed-Interconnect) VAXcluster system:

- Install or upgrade to the latest version of the VMS operating system.
- Answer questions in the installation/upgrade procedure to configure the boot server.
- Allow Ethernet VAXcluster communications.
- Execute SYS$MANAGER:CLUSTER_CONFIG.COM on this member to add CI systems.
- Enable some or all CI members as boot servers.
- Execute SYS$MANAGER:CLUSTER_CONFIG.COM on the boot server to add satellites.
- Execute SYS$MANAGER:CLUSTER_CONFIG.COM on the boot server to create an additional system disk, if desired.
- Continue setting up cluster devices, disks, tapes, etc.

Figure 5–5 A Mixed-Interconnect VAXcluster System with a Phase I Shadow Set as a System Disk



Satellite System Disks on an HSC Controller

Putting satellite system disks on an HSC or DSSI bus in a mixed-interconnect cluster:

- Gives the most flexibility
- Provides highest availability
- Lets any CI or DSSI member be enabled as a boot server
- Lets any CI or DSSI member be selected to serve disks to satellites

Adding Ethernet Members to an Existing CI Cluster

- Enable cluster communications over the Ethernet on all VAX nodes
  - Log in as system manager on each VAX node.
  - Invoke CLUSTER_CONFIG using the CHANGE function to enable the Ethernet for cluster communications.
- Enable one or more nodes as boot servers, as follows:
  - Log in as system manager on each VAX node.
  - Execute the CHANGE function to enable one or more nodes as boot servers, and to change ALLOCLASS if it is necessary.
- Use CLUSTER_CONFIG ADD to add satellite roots to a CI system disk, if desired.
- If the number of votes has changed:
  - Edit MODPARAMS.DAT to set a new appropriate value for EXPECTED_VOTES if the number of votes has changed.
  - Execute AUTOGEN on all members.
  - Shut down and reboot the cluster.

Adding a CI Interconnect to an Existing Local Area Cluster

If the boot server is a CI capable node:

- Log in as system manager on the boot server.
- Perform an image backup operation to back up the current system disk to a disk on an HSC.
- Modify the system's default bootstrap command procedure to boot the system from the HSC disk.
- Shut down the cluster.
  (Shut down the satellites first, then shut down the boot server.)
- Boot the boot server from the newly created system disk on the HSC.
- Reboot the satellites.

If your current boot server is not a CI capable node:

- Shut down the old local area cluster.
  (Shut down the satellites first, then shut down the boot server.)
- Install the VMS operating system on the new CI connected VAX node's HSC system disk.
  - When the installation procedure asks if you want to enable the Ethernet for cluster communications, answer YES.
- When the installation completes, log in as system manager.
  - Install PAKs.
  - Configure and start the DECnet VAX network.
  - Execute the CLUSTER_CONFIG CHANGE function to enable the node as a boot server.
  - Execute the CLUSTER_CONFIG ADD function to add the former local area cluster members (including the former boot server) as satellites on the new HSC system disk.

NOTE

This procedure places the new system disk for the resulting mixed-interconnect cluster on an HSC disk. This is not required, but highly recommended.

Merging an Existing Local Area Cluster and an Existing CI Cluster

The following procedure describes the minimum amount of work needed to merge an existing local area cluster and an existing CI only cluster into a single cluster.

- Execute CLUSTER_CONFIG CHANGE on all CI members to allow cluster communication over the Ethernet.
  - The procedure asks for the cluster group number and password.
  - Give the same number and password as are being used by the local area cluster.
- If the boot server from the old local area cluster is CI capable, then it should be connected to the star coupler. (Ask your Customer Service representative to do this for you.)
- To enhance availability, the system disks for all satellites should be moved to HSC disks, if at all possible.
- Determine the appropriate VOTE configuration for the new mixed-interconnect cluster.
- Enter the total number of VOTES as the value of the EXPECTED_VOTES parameter in MODPARAMS.DAT on each member.
- Reboot all old CI members.
- Boot the old local area cluster boot server.
- Boot the Ethernet satellites.

Table 5–3 CLUSTER_CONFIG.COM CHANGE Options

Option	Operation Performed
Enable the local system as a disk server.	Load the MSCP server by setting, in MODPARAMS.DAT, the value of the MSCP_LOAD parameter to 1 and setting an appropriate value for the MSCP_SERVE_ALL parameter.
Disable the local system as a disk server.	Set MSCP_LOAD to 0.
Enable the local system as a boot server.	If you are setting up a local area or mixed-interconnect cluster, you must execute this operation once before you attempt to add nodes to the cluster. You thereby enable DECnet MOP service for the Ethernet adapter circuit that the node will use to service down-line load requests from satellites. When you enable the node as a boot server, it automatically becomes a disk server (if it is not one already), because it must serve its system disk to satellites.
Set WINDOW_SYSTEM=0.	
Disable the local system as a boot server.	Disable DECnet MOP service for the node's Ethernet adapter circuit.
Enable the Ethernet for cluster communications on the local system.	Load the VAXport driver PEDRIVER by setting the value of the NISCS_LOAD_PEA0 parameter to 1 in MODPARAMS.DAT. Create the cluster security database file, SYS$SYSTEM:[SYSEXE]CLUSTER_AUTHORIZE.DAT, on the local system's system disk.
Disable the Ethernet for cluster communications on the local system.	Set NISCS_LOAD_PEA0 to 0.
Enable a quorum disk on the local system.	Set, in MODPARAMS.DAT, an appropriate value for the SYSGEN parameter DISK_QUORUM; set the value of QDSKVOTES to 1 (default value).
Disable a quorum disk on the local system.	Set, in MODPARAMS.DAT, a blank value for the SYSGEN parameter DISK_QUORUM; set the value of QDSKVOTES to 1 (default value).
Change the local system's allocation class value.	Set a value for the node's ALLOCLASS parameter in MODPARAMS.DAT.
Change a satellite's Ethernet hardware address.	Change a satellite's hardware address, in the event that its Ethernet device should need replacement. Both the permanent and volatile network databases, and NETNODE_UPDATE.COM, are updated on the local system. **You must execute this operation on any node enabled as a boot server for the satellite.**

Example 5–6 Sample Interactive CLUSTER_CONFIG.COM Session to Enable the Local System as a Disk Server

```
$ @CLUSTER_CONFIG
                    Cluster Configuration Procedure

    Use CLUSTER_CONFIG to set up or change a VAXcluster configuration.
    To ensure that you have the required privileges, invoke this procedure
    from the system manager's account.

    Enter ? for help at any prompt.

            1. ADD a node to the cluster.
            2. REMOVE a node from the cluster.
            3. CHANGE a cluster node's characteristics.
            4. CREATE a second system disk for BARNUM.

    Enter choice [1]: 3

    CHANGE Menu

       1. Enable BARNUM as a disk server.
       2. Disable BARNUM as a disk server.
       3. Enable BARNUM as a boot server.
       4. Disable BARNUM as a boot server.
       5. Enable Ethernet for cluster communications on BARNUM.
       6. Disable Ethernet for cluster communications on BARNUM.
       7. Enable a quorum disk on BARNUM.
       8. Disable a quorum disk on BARNUM.
       9. Change BARNUM's ALLOCLASS value.
      10. Change a satellite's hardware address.

    Enter choice [1]: RETURN
Will BARNUM serve HSC disks [Y]? RETURN
Enter a value for BARNUM's ALLOCLASS parameter: 2
The configuration procedure has completed successfully.

    BARNUM has been enabled as a disk server. MSCP_LOAD has been
    set to 1 in MODPARAMS.DAT. Please execute AUTOGEN to reboot BARNUM:

            $ @SYS$UPDATE:AUTOGEN GETDATA REBOOT

    If you have changed BARNUM's ALLOCLASS value, you must reconfigure the
    cluster, using the procedure described in the VMS VAXcluster Manual.
```

Example 5-7 Sample Interactive CLUSTER_CONFIG.COM Session to Change the Local System's ALLOCLASS Value

```
$ @CLUSTER_CONFIG

                 Cluster Configuration Procedure

    Use CLUSTER_CONFIG to set up or change a VAXcluster configuration.
    To ensure that you have the required privileges, invoke this procedure
    from the system manager's account.

    Enter ? for help at any prompt.

            1. ADD a node to the cluster.
            2. REMOVE a node from the cluster.
            3. CHANGE a cluster node's characteristics.
            4. CREATE a second system disk for BARNUM.

    Enter choice [1]: 3

    CHANGE Menu

       1. Enable BARNUM as a disk server.
       2. Disable BARNUM as a disk server.
       3. Enable BARNUM as a boot server.
       4. Disable BARNUM as a boot server.
       5. Enable Ethernet for cluster communications on BARNUM.
       6. Disable Ethernet for cluster communications on BARNUM.
       7. Enable a quorum disk on BARNUM.
       8. Disable a quorum disk on BARNUM.
       9. Change BARNUM's ALLOCLASS value.
      10. Change a satellite's hardware address.

    Enter choice [1]: 9

Enter a value for BARNUM's ALLOCLASS parameter [2]: 1
The configuration procedure has completed successfully

    If you have changed BARNUM's ALLOCLASS value, you must reconfigure the
    cluster, using the procedure described in the VMS VAXcluster Manual.
```

Example 5–8 Sample Interactive CLUSTER_CONFIG.COM Session to Enable the Local System as a Boot Server

```
$ @CLUSTER_CONFIG
                Cluster Configuration Procedure

    Use CLUSTER_CONFIG to set up or change a VAXcluster configuration.
    To ensure that you have the required privileges, invoke this procedure
    from the system manager's account.

    Enter ? for help at any prompt.

            1. ADD a node to the cluster.
            2. REMOVE a node from the cluster.
            3. CHANGE a cluster node's characteristics.
            4. CREATE a second system disk for BARNUM.

    Enter choice [1]: 3

    CHANGE Menu

         1. Enable BARNUM as a disk server.
         2. Disable BARNUM as a disk server.
         3. Enable BARNUM as a boot server.
         4. Disable BARNUM as a boot server.
         5. Enable Ethernet for cluster communications on BARNUM.
         6. Disable Ethernet for cluster communications on BARNUM.
         7. Enable a quorum disk on BARNUM.
         8. Disable a quorum disk on BARNUM.
         9. Change BARNUM's ALLOCLASS value.
        10. Change a satellite's hardware address.

    Enter choice [1]: 3

Verifying circuits in network database...
Updating permanent network database...

    In order to enable or disable DECnet MOP service in the volatile
    network database, DECnet traffic must be interrupted temporarily.

    Do you want to proceed [Y]? [RETURN]

Enter a value for BARNUM's ALLOCLASS parameter [1]: [RETURN]
The configuration procedure has completed successfully.

    BARNUM has been enabled as a boot server. Disk serving and
    Ethernet capabilities are enabled automatically. If BARNUM was
    not previously set up as a disk server, please execute AUTOGEN to
    reboot BARNUM:

            $ @SYS$UPDATE:AUTOGEN GETDATA REBOOT

    If you have changed BARNUM's ALLOCLASS value, you must reconfigure the
    cluster, using the procedure described in the VMS VAXcluster Manual.
```

Example 5–9 Sample Interactive CLUSTER_CONFIG.COM Session to Change a Satellite's Hardware Address

```
$ @CLUSTER_CONFIG

                    Cluster Configuration Procedure

    Use CLUSTER_CONFIG to set up or change a VAXcluster configuration.
    To ensure that you have the required privileges, invoke this procedure
    from the system manager's account.

    Enter ? for help at any prompt.

            1. ADD a node to the cluster.
            2. REMOVE a node from the cluster.
            3. CHANGE a cluster node's characteristics.
            4. CREATE a second system disk for BARNUM.

    Enter choice [1]: 3

    CHANGE Menu

       1. Enable BARNUM as a disk server.
       2. Disable BARNUM as a disk server.
       3. Enable BARNUM as a boot server.
       4. Disable BARNUM as a boot server.
       5. Enable Ethernet for cluster communications on BARNUM.
       6. Disable Ethernet for cluster communications on BARNUM.
       7. Enable a quorum disk on BARNUM.
       8. Disable a quorum disk on BARNUM.
       9. Change BARNUM's ALLOCLASS value.
      10. Change a satellite's hardware address.

    Enter choice [1]: 10

What is the node's DECnet node name? TIGER
What is the new hardware address [08-00-86-21-34-76]? 08-00-3B-05-37-78
Updating network database...
The configuration procedure has completed successfully.
```

BOOT COMMAND PROCEDURES

Processor Startup

- Cluster members can be made to automatically boot into the cluster.
- Satellite nodes will request a down-line load over the Ethernet.
- VAX members may need boot command procedures edited in order for correct devices and disks to be referenced.
- **CI and DSSI nodes:**
  - — Boot from a disk connected to an HSC or from an ISE unit

    (or)
  - — Boot from a local disk
- Specify HSC CI node numbers (or two node numbers if the system disk is dual-ported).
- Specify the unit number of the system disk (HSC disk).
- If the system disk is shadowed:
  - — Phase I (controller-based): Specify the virtual unit number of the shadow set and the unit number of one of the disks in the shadow set.
  - — Phase II (host-based): Specify the physical unit number of a locally accessible member (local, HSC, or DSSI disk) of the shadow set. SYSGEN parameters will specify that the system disk is shadowed and provide the virtual unit number of the shadow set.
- Specify the system root on the system disk.
- Specify other processor-dependent information.
- Booting from a DSSI disk is just like booting from any other kind of local disk.
  - — Using KFQSA: `>>>BOOT DUAn`
  - — Using EDA640:`>>>BOOT DIAn`
  - — Add `/R5:r000000` if booting from root r

MicroVAX and VAXstation satellites can be configured to automatically boot into the cluster or to boot standalone from a system on their local disks.

CI Boot Command Procedures

Create boot procedures on your console medium to simplify later booting. Use the customizable template boot file CIBOO.CMD to create the following, specific boot files:

- DEFBOO.CMD performs a nonstop boot
- GENBOO.CMD performs an interactive boot
- CI.CMD allows any type of boot from any root

To create the procedures:

- Copy the customizable template CIBOO.CMD from the console medium

```
RUN SYS$SYSTEM:SYSGEN
SYSGEN> CONNECT CONSOLE
SYSGEN> EXIT
EXCHANGE COPY CSA1:CIBOO.CMD *.*/LOG
```

- Copy CIBOO.CMD to DEFBOO.CMD, GENBOO.CMD, and CI.CMD
- Edit DEFBOO.CMD, GENBOO.CMD, and CI.CMD to perform their intended functions
- Copy the new DEFBOO.CMD, GENBOO.CMD, and CI.CMD back to the console medium

```
EXCHANGE COPY DEFBOO.CMD,GENBOO.CMD,CI.CMD  CSA1:*.*/LOG
DISMOUNT CSA1:
```

Common Boot Command Procedure Attributes

Register meaning in CI boot command procedures:

- Register 0: VAX CI port device code
- Register 1: processor bus information (VAXBI device number on VAXBI processors)
- Register 2: HSC node numbers of the form xxyy (hexadecimal)
  - xx and yy are the HSC CI node numbers of any HSC ported to the disk (assuming the disk is dual-ported between the HSC). If one of these numbers is 0, it must be in the low-order byte. Indien ØØ, dan als laatste opgeven bv. Ø1ØØ
- Register 3: disk unit number and virtual unit number for shadow sets
  - Convert unit numbers to hexadecimal
  - Unit number of disk (for example, DUA4: would be indicated as `D/G 3 4`)
  - When booting from a Phase I (controller-based) shadow set:

    Put unit number of **virtual** unit number in the **high** word (bits <30:16>)

    Put the physical unit number of one of the members of the shadow set into the **low** word (bits <15:0>)

    CAUTION

    All VAXcluster members using Phase I (controller-based) shadowed system disks must specify the same physical unit in their boot procedures or different members might try to build the shadow set based on different disks, losing system crash information.

    Set the sign bit (bit 31) to indicate you are using a shadow set virtual unit (the sign bit is set when the high-order hexadecimal digit is 8)

    For example, Physical unit: 1, Virtual unit: 12 (decimal) = `D/G 3 800C0001`

    is een shadow set 12 #

- Register 4: not used
- Register 5: system root on the system disk and method of boot, in the form (r000000i)
  - r is 0<=r<=F (hex)
  - i=0 for non-stop boot
  - i=1 for conversational boot
- For example, root=3, i=0 for a nonstop boot from [SYS3], `D/G 5 30000000`
- For example, root=3, i=1 for a conversational boot, `D/G 5 30000001`

For more information, see your processor installation guide or owner's manual.

Booting from an HSC Disk

You can boot by either:

- Typing a series of console commands
- Executing a boot command procedure

Either method must deposit values in certain registers:
(using the deposit commands appropriate to the processor)

- Deposit HSC node numbers (hexadecimal) in R2

  `DEPOSIT R2 xxyy`

- Deposit the system disk unit number (hexadecimal) in R3

  `DEPOSIT R3 uu`

  (or)

  `DEPOSIT R3 8vvvuuuu` for a Phase I shadowed system disk

- Deposit boot flags in R5

  `DEPOSIT R5 r000000i`

  — **r** is the root number (hexadecimal)

  — **i** is 0 for a nonstop boot, 1 for a conversational boot

- Enter the command

  `>>>@ CI.CMD`

Initial Boot from a New Root

CLUSTER_CONFIG sets up all the necessary activity for the initial boot.

This includes:

- Using a special-purpose STARTUP file (STARTUP1.COM)
- Setting SYSGEN parameters
- For W-Kit processors, executes LMF$CONFIG.COM, creating the VMS license for the member
  - MicroVAX II system

    LMF$CONFIG.COM asks how many users on that system.

    You must supply the answer before the boot can complete.
- Rebooting the system with normal startup procedures once initial parameters have been set

CONFIGURING SYSTEM DISKS

Files that Must Remain on the System Disk

The sizes of these files and the amount of available space on a system disk determine how many systems can boot from a single volume.

- VMS system executable files
- Startup files
- Error logs
- SYSDUMP.DMP for each member that boots from the system disk
- DCL tables
- Network databases

Files that can be Moved Off the System Disk

These files do not need to be on the system disk. They are located by logical names on each member.

Table 5–4 System Files and Their Logical Names

File	Logical Name	System Use
SYSUAF.DAT	SYSUAF	System user authorization file
NETPROXY.DAT	NETPROXY	System network proxy file
RIGHTSLIST.DAT	RIGHTSLIST	System UIC name identifiers
VMSMAIL_PROFILE.DATA	VMSMAIL_PROFILE	VMSMAIL data file
LMF$LICENSE.LDB	LMF$LICENSE	LMF cluster common database
JBCSYSQUE.DAT	JBCSYSQUE	Job controller file
ACCOUNTNG.DAT	ACCOUNTNG	Accounting file

System Memory Size and SYSDUMP.DMP

When SYSGEN parameter DUMPBUG is set to 1, SYSDUMP.DMP is created upon system failure.

- The size of SYSDUMP.DMP can limit the number of nodes that can boot from a single system disk.
- SYSGEN parameter DUMPSTYLE set to 1 will cause only selected portions of the dump file to be written.
- Setting this parameter to 1 saves large amounts of disk space.

Table 5–5 Possible Values for SYSGEN Parameter DUMPSTYLE

Value	Meaning
0	Entire contents of physical memory will be written to the dump file. (This is the default.)
1	Selective portions of memory will be written to the dump file as space permits.

SETTING UP THE VAXcluster OPERATING ENVIRONMENT

To set up the operating environment, create or modify the following:

- SYSGEN parameters
- SYSUAF, NETPROXY, RIGHTSLIST, and VMSMAIL_PROFILE and LMF$LICENSE databases
- Cluster-common and node-specific startup command procedures
- Queue management command procedures
- Disk management command procedures
- DECnet startup command procedures

Using MODPARAMS.DAT to Modify SYSGEN Parameters

- Some SYSGEN parameters that are assigned by CLUSTER_CONFIG, but which might need occasional changes:
  - — VOTES
  - — EXPECTED_VOTES
  - — VAXCLUSTER
  - — MCSP_LOAD
  - — MSCP_SERVE_ALL
  - — ALLOCLASS
- To display VAXcluster-related SYSGEN parameters:

```
$ RUN SYS$SYSTEM:SYSGEN
SYSGEN> SHOW /CLUSTER
SYSGEN> SHOW /SCS
SYSGEN> CTRL/Z
```

- To modify SYSGEN parameters:
  - — Edit SYS$SPECIFIC:[SYSEXE]MODPARAMS.DAT
  - — Execute SYS$UPDATE:AUTOGEN.COM
- Advantages of using AUTOGEN rather than SYSGEN:
  - — AUTOGEN uses a feedback mechanism to set parameter values based on your system's workload
  - — AUTOGEN adjusts other parameters to reflect your changes
  - — MODPARAMS.DAT, in effect, maintains a record of changes that are not lost during VMS updates

Example 5–10 Sample MODPARAMS.DAT After a VMS Upgrade (Sheet 1 of 2)

```
!++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
! SYS$SYSDEVICE:[SYS0.SYSEXE]MODPARAMS.DAT
!
! This is a new file created by the VMS upgrade procedure.  This file
! was built by using the data found in the following file(s) previously
! used by this system:
!
!  SYS$SYSDEVICE:[SYS0.SYSEXE]VAXVMSSYS.PAR
!  SYS$SYSDEVICE:[SYS0.SYSEXE]PARAMS.DAT
!  SYS$SYSDEVICE:[SYS0.SYSEXE]MODPARAMS.DAT
!
! Your old file(s) have been renamed to:
!
!  SYS$SYSDEVICE:[SYS0.SYSEXE]VAXVMSSYS.PAR_OLD
!  SYS$SYSDEVICE:[SYS0.SYSEXE]MODPARAMS.DAT_OLD
!
! In order to provide a bootable environment a new
!
!  SYS$SYSDEVICE:[SYS0.SYSEXE]VAXVMSSYS.PAR
!
! had to be provided. Your old file is not compatible with this release.
! Previous parameters found to be larger than the new defaults were
! retained.
!
! Please check the following sections of this file for an indication of
! what files were used, and in what sequence, in providing additional
! parameter data.  Please review and edit this file for possible
! duplications, additions and deletions you wish to make.
!
!----------------------------------------------------------------------------
!****************************************************************************
! This section contains any VAXcluster parameters found in
! SYS$SYSDEVICE:[SYS0.SYSEXE]VAXVMSSYS.PAR with values different than
! the defaults.
!
SCSSYSTEMID=19880
SCSNODE="HORSE  "
VAXCLUSTER=2
EXPECTED_VOTES=1
VOTES=1
RECNXINTERVAL=20
DISK_QUORUM="                            "
QDSKVOTES=1
QDSKINTERVAL=20
ALLOCLASS=0
LOCKDIRWT=5
NISCS_CONV_BOOT=1
NISCS_LOAD_PEA0=1
NISCS_PORT_SERV=0
MSCP_LOAD=1
MSCP_SERVE_ALL=1
```

Example 5–10 Sample MODPARAMS.DAT After a VMS Upgrade (Sheet 2 of 2)

```
!**************************************************************************
! This section contains any OLDSITE parameters found in
! SYS$SYSDEVICE:[SYS0.SYSEXE]PARAMS.DAT
!
!**************************************************************************
! This section contains any parameters found in
! SYS$SYSDEVICE:[SYS0.SYSEXE]MODPARAMS.DAT
!
SCSSYSTEMID=19880
SCSNODE="HORSE  "
VAXCLUSTER=2
EXPECTED_VOTES=1
VOTES=1
RECNXINTERVAL=20
DISK_QUORUM="                "
QDSKVOTES=1
QDSKINTERVAL=20
ALLOCLASS=0
LOCKDIRWT=5
NISCS_CONV_BOOT=1
NISCS_LOAD_PEA0=1
NISCS_PORT_SERV=0
SWAPFILE=15000
MSCP_SERVE_ALL=1
MSCP_LOAD=1
GBLPAGES=20000
GBLSECTIONS=500
PAGEFILE=30000
```

Mail and User Authorization Databases

VMS automatically creates the following files in SYS$SPECIFIC:[SYSEXE]:

- SYSUAF.DAT
- NETPROXY.DAT
- RIGHTSLIST.DAT
- VMSMAIL_PROFILE.DATA

In a common environment VAXcluster system, these files should be moved to SYS$COMMON:[SYSEXE].

To create VMSMAIL_PROFILE.DATA in SYS$COMMON:

- Define the logical name VMSMAIL_PROFILE before any user enters MAIL

```
$ DEFINE/SYSTEM/EXEC VMSMAIL_PROFILE SYS$COMMON:[SYSEXE]VMSMAIL_PROFILE
```

- Invoke Mail, which creates the VMSMAIL_PROFILE.DATA database

```
$ MAIL
```

- You do not need the logical name VMSMAIL_PROFILE after Mail has created VMSMAIL_PROFILE.DATA

You can set the logical name MAIL$SYSTEM_FLAGS to:

- 1 to disable the use of the DECnet network to send Mail to another node in the VAXcluster system
- 2 to notify a user of new Mail on every node where that user is logged in
- 4 to include the time in a Mail notification message
- 7 to do all three of the above

```
$ DEFINE/SYSTEM/EXEC MAIL$SYSTEM_FLAG 7
```

For ease of management, a common environment VAXcluster system should have one common copy of each of these files, even if there are several system disks:

- SYSUAF.DAT
- RIGHTSLIST.DAT
- NETPROXY.DAT
- VMSMAIL_PROFILE.DATA

Management tasks include:

- Placing logical names in the SYLOGICALS.COM startup command procedure for each node

```
$ DEFINE/SYSTEM/EXEC MAN_DSK $1$DUA1:
$ DEFINE/SYSTEM/EXEC SYSUAF  MAN_DSK:[CLUSMAN]SYSUAF.DAT
$ DEFINE/SYSTEM/EXEC RIGHTSLIST MAN_DSK:[CLUSMAN]RIGHTSLIST.DAT
$ DEFINE/SYSTEM/EXEC NETPROXY  MAN_DSK:[CLUSMAN]NETPROXY.DAT
$ DEFINE/SYSTEM/EXEC VMSMAIL_PROFILE  MAN_DSK:[CLUSMAN]VMSMAIL_PROFILE.DATA
```

- Merging like files from existing VMS systems:
  - Obtain a listing of the records in each file
  - Eliminate duplicates and conflicts
  - Merge all records into a single file
  - Place the resulting file on a cluster-accessible disk
- Using the CONVERT utility to merge:
  - SYSUAF.DATs
  - NETPROXY.DATs
  - VMSMAIL_PROFILE.DATAs
- Making user disks available cluster-wide

An Example of Merging SYSUAF.DAT Files from Existing Systems

- Create SYSUAF.LIS for each node

```
$ SET DEFAULT SYS$SYSTEM
$ RUN AUTHORIZE
UAF> LIST
UAF> EXIT
```

- Copy SYSUAF.DAT on HORSE to HORSE_SYSUAF.DAT
- Copy SYSUAF.DAT on BAILEY to BAILEY_SYSUAF.DAT
- Find duplicate records in HORSE and BAILEY authorization files

1 van de dubbele namen komt hier in.

```
$ CONVERT/EXCEPTIONS=EXC.DAT -
_$ HORSE_SYSUAF.DAT,BAILEY_SYSUAF.DAT COMBINED_SYSUAF.DAT
```

- Compare the SYSUAF.LIS from HORSE to the SYSUAF.LIS from BAILEY. Look for:
  - — Duplicate user names by examining EXC.DAT
  - — Different user names with the same UIC and out-of-date UIC records by examining SYSUAF.LIS
- Use the AUTHORIZE utility to remove or modify SYSUAF records for each file
- Merge prepared SYSUAF files

```
$ CONVERT/EXCEPTIONS=EXC.DAT -
_$ HORSE_SYSUAF.DAT,BAILEY_SYSUAF.DAT COMBINED_SYSUAF.DAT
```

- Copy the merged SYSUAF file to the appropriate location

```
$ COPY COMBINED_SYSUAF.DAT $1$DUA1:[CLUSMAN]SYSUAF.DAT
```

An Example of Creating the Rights Database

- Create RIGHTSLIST.LIS for each node. The following procedure shows all rightslist identifiers.

```
$ SET DEFAULT SYS$SYSTEM
$ RUN AUTHORIZE
UAF> LIST/IDENTIFIER *
UAF> EXIT
```

- Create RIGHTSLIST.LIS showing which users hold which identifiers.

```
UAF> LIST/RIGHTS/USER=*
```

- Create the RIGHTSLIST logical name.

```
$ DEFINE/SYSTEM/EXEC RIGHTSLIST $1$DUA1:[CLUSMAN]RIGHTSLIST.DAT
```

- Create 1DUA1:[CLUSMAN]RIGHTSLIST.DAT

```
$ RUN AUTHORIZE
UAF> CREATE/RIGHTS
UAF> ADD/IDENTIFIER/USER=*
UAF> EXIT
```

- Examine each RIGHTSLIST.LIS for identifiers created by the system manager.
- Add these identifiers to RIGHTSLIST.DAT and grant them to the appropriate users.

```
UAF> ADD/IDENTIFIER id-name
UAF> GRANT/ID id-name user-spec
```

An Example of Merging VMSMAIL_PROFILE.DATA Files from Existing Systems

- Copy VMSMAIL_PROFILE.DATA on HORSE to HORSEMAIL.DAT and VMSMAIL_PROFILE.DATA on BAILEY to BAILEYMAIL.DAT

  ```
  $ CONVERT/SHARE VMSMAIL_PROFILE.DATA HORSEMAIL.DAT
  ```

  (or)

  ```
  $ COPY VMSMAIL_PROFILE.DATA HORSEMAIL.DAT
  ```

- Merge the Mail files.

  ```
  $ CONVERT/EXCEPTIONS=EXC.DAT -
  _$ HORSEMAIL.DAT,BAILEYMAIL.DAT VMSMAIL_PROFILE.DATA
  ```

- Examine the exceptions file and decide how to handle duplicate records.
  - Use the REMOVE command in Mail utility to remove obsolete records.
- Merge the Mail files again (as in step 2) to produce a combined file.
- Copy the new combined VMSMAIL_PROFILE.DATA to the appropriate location.

  ```
  $ COPY VMSMAIL_PROFILE.DATA $1$DUA1:[CLUSMAN]VMSMAIL_PROFILE.DATA
  ```

Multiple Environment VAXcluster Systems

In a multiple environment cluster, each system can have its own:

- SYSUAF.DAT
- RIGHTSLIST.DAT (see below)
- NETPROXY.DAT
- VMSMAIL_PROFILE.DATA

Management tasks include:

- Making separate SYSUAF.DAT files compatible
  - No two users should have the same user name
  - No two users should have the same UIC (without good reason)
- Coordinating RIGHTSLIST identifiers
  - Users on different systems should have different identifiers, unless you want to give them access to the same objects.
  - Even in multiple environment clusters, there should be only one RIGHSLIST.DAT.

    Coordinating Access Control Lists (ACLs) between multiple RIGHTSLISTs is confusing, time-consuming, and, ultimately, unmanageable.

    RIGHTSLIST.DAT should always be a common file because disks are normally shared, and the most common use of Rights/Identifiers is in access control lists on files.
- Changing proxy accounts for DECnet access
  - To list users, identifiers, and proxies:

```
$ SET DEFAULT SYS$SYSTEM
$ RUN AUTHORIZE
UAF> LIST
UAF> LIST/IDENTIFIER *
UAF> LIST/PROXY
```

Startup Command Procedures

- Configure devices with commands such as:
  - SET DEVICE/DUAL_PORT
  - SET PRINTER
  - SET TERMINAL
- Start local device and batch queues
  - Mount disks
  - Define logical names
  - Start up layered products
  - Start DECnet software and the LAT driver

Three ways to set up common command procedures:

1. Have each node execute the same command procedure:

   SYS$COMMON:[SYSMGR]SYSTARTUP_V5.COM

   - Use conditional statements to branch to node-specific parts of the procedure

     ```
     IF F$GETSYI("NODENAME") .EQS. "VAXA"
     ```

2. Have SYS$SPECIFIC:[SYSMGR]SYSTARTUP_V5.COM call

   SYS$COMMON:[SYSMGR]COMMON_STARTUP.COM

3. Have SYS$COMMON:[SYSMGR]SYSTARTUP_V5.COM call member specific startup command procedures located in

   SYS$SPECIFIC:[SYSMGR]

If a common environment is not desired, a member in the cluster could simply have a private command procedure SYS$SPECIFIC:[SYSMGR]SYSTARTUP_V5.COM.

Cluster-Common Startup Command Procedure that Executes Specific Procedures

Example 5–11 Sample SYSTARTUP_V5.COM

```
$! SYSTARTUP_V5.COM
$! Site-specific startup command procedure for cluster
$ SET NOON
$!
$! Execute node-specific procedure to mount disks
$ @SYS$SPECIFIC:[SYSMGR]MOUNT.COM
$!
$! Execute node-specific terminal set-up procedure
$ @SYS$SPECIFIC:[SYSMGR]TERMINALS
$!
$! Execute node-specific queue startup procedure
$ @SYS$SPECIFIC:[SYSMGR]STARTQUE
$!
$! Install the same images on each node
$ INSTALL
CREATE FORTRAN/SHARED/OPEN
CREATE MACRO32/SHARED/OPEN
$!
$ PURGE/KEEP=2 SYS$SPECIFIC:[SYSMGR]OPERATOR.LOG
$!
$! Start up DECnet -- in node-specific queue!
$ NODE = F$GETSYI("NODENAME")
$ SUBMIT/NOPRINT/QUEUE='NODE'_BATCH -
  SYS$SPECIFIC:[SYSMGR]STARTNET, SYS$SPECIFIC:[SYSMGR]LTLOAD
$!
$! Start up the same layered products on each node
$ @SYS$STARTUP:SPM$STARTUP     !Startup SPM
$ @SYS$STARTUP:VPA$STARTUP     !Startup VPA
$ EXIT
```

Setting up Cluster Queues

JBCSYSQUE.DAT:

- Stores cluster-wide queue information on a cluster-available disk
- Makes queue names unique to avoid conflict in the cluster-wide queue file

Setting Up Cluster-Wide Batch and Print Queues

- Place the following code in the startup command procedure for each node:

  ```
  $ START/QUEUE/MANAGER SYS$COMMON:[SYSEXE]JBCSYSQUE.DAT
  ```

  (or)

  ```
  $ START/QUEUE/MANAGER some_common_disk:[directory]JBCSYSQUE.DAT
  ```

- Initialize queues from each node where they will be used
- Start each queue on the node where its print or CPU resource is located
- Include the /ON=resource_name qualifier to specify the exact name of the printer or CPU as needed
- Resource_name is of the form:
  - **node::** for a batch queue
  - **node::device:** for a print queue
- Edit access control lists to restrict access to queues as desired:

  ```
  $ SET QUEUE/PROTECTION=WORLD LN03_PRINT
  $ SET ACL/OBJECT_TYPE=QUEUE/ACL=(IDENTIFIER=SECRETARIES, ACCESS=WRITE) -
  _$ LN03_PRINT
  ```

Examples of Creating Cluster-Wide Generic Print and Batch Queues

Example 5–12 Queue Creation Commands on BARNUM

```
$ INITIALIZE/QUEUE/ON=BARNUM::LPA0:/START   BARNUM_PRINT
$ INITIALIZE/QUEUE/ON=BAILEY::LPA0:         BAILEY_PRINT
$ INITIALIZE/QUEUE/GENERIC=(BARNUM_PRINT,BAILEY_PRINT) -
  /START CLUSTER_PRINT
$ DEFINE/SYSTEM SYS$PRINT CLUSTER_PRINT
```

Example 5–13 Queue Creation Commands on BAILEY

```
$ INITIALIZE/QUEUE/ON=BARNUM::LPA0:         BARNUM_PRINT
$ INITIALIZE/QUEUE/ON=BAILEY::LPA0:/START   BAILEY_PRINT
$ INITIALIZE/QUEUE/GENERIC=(BARNUM_PRINT,BAILEY_PRINT) -
  /START CLUSTER_PRINT
$ DEFINE/SYSTEM SYS$PRINT CLUSTER_PRINT
```

Example 5–14 Queue Creation Commands on All Other Cluster Members

```
$ INITIALIZE/QUEUE/ON=BARNUM::LPA0:         BARNUM_PRINT
$ INITIALIZE/QUEUE/ON=BAILEY::LPA0:         BAILEY_PRINT
$ INITIALIZE/QUEUE/GENERIC=(BARNUM_PRINT,BAILEY_PRINT) -
  /START CLUSTER_PRINT
$ DEFINE/SYSTEM SYS$PRINT CLUSTER_PRINT
```

Example 5–15 SHOW QUEUE Output on Any Cluster Member

```
$ SHOW QUEUE/DEVICE/FULL
Printer queue BARNUM_PRINT, on BARNUM::LPA0:
/BASE_PRIORITY=4 /DEFAULT=(FEED) /FORM=DEFAULT
/OWNER=[SYSTEM] /PROTECTION=(S:E,O:D,G:R,W:W)

Printer queue BAILEY_PRINT, on BAILEY::LPA0:
/BASE_PRIORITY=4 /DEFAULT=(FEED) /FORM=DEFAULT
/OWNER=[SYSTEM] /PROTECTION=(S:E,O:D,G:R,W:W)

Generic printer queue CLUSTER_PRINT
/GENERIC=(BARNUM_PRINT,BAILEY_PRINT) /OWNER=[SYSTEM]
/PROTECTION=(S:E,O:D,G:R,W:W)
```

Creating a Cluster-Wide Generic Batch Queue

Example 5-16 Creating and Displaying Cluster-Wide Batch Queues

On LION:

```
$ INITIALIZE/QUEUE/BATCH/ON=LION::/START   LION_BATCH
$ INITIALIZE/QUEUE/BATCH/ON=TIGER::        TIGER_BATCH
$ INITIALIZE/QUEUE/BATCH/ON=BEAR::         BEAR_BATCH
$ INITIALIZE/QUEUE/BATCH/ON=HORSE::        HORSE_BATCH
$ INITIALIZE/QUEUE/BATCH/ON=BARNUM::       BARNUM_BATCH
$ INITIALIZE/QUEUE/BATCH/ON=RNGLNG::       RNGLNG_BATCH
$ INITIALIZE/QUEUE/BATCH/ON=BAILEY::       BAILEY_BATCH
$ INITIALIZE/QUEUE/BATCH/GENERIC=(LION_BATCH,TIGER_BATCH,BEAR_BATCH,-
  HORSE_BATCH,BARNUM_BATCH,RNGLNG_BATCH,BAILEY_BATCH)/START  SYS$BATCH
```

On BAILEY:

```
$ INITIALIZE/QUEUE/BATCH/ON=LION::         LION_BATCH
$ INITIALIZE/QUEUE/BATCH/ON=TIGER::        TIGER_BATCH
$ INITIALIZE/QUEUE/BATCH/ON=BEAR::         BEAR_BATCH
$ INITIALIZE/QUEUE/BATCH/ON=HORSE::        HORSE_BATCH
$ INITIALIZE/QUEUE/BATCH/ON=BARNUM::       BARNUM_BATCH
$ INITIALIZE/QUEUE/BATCH/ON=RNGLNG::       RNGLNG_BATCH
$ INITIALIZE/QUEUE/BATCH/ON=BAILEY::/START BAILEY_BATCH
$ INITIALIZE/QUEUE/BATCH/GENERIC=(LION_BATCH,TIGER_BATCH,BEAR_BATCH,-
  HORSE_BATCH,BARNUM_BATCH,RNGLNG_BATCH,BAILEY_BATCH)/START  SYS$BATCH
```

On any cluster member:

```
$ SHOW QUEUE/BATCH/FULL
Batch queue HORSE_BATCH, on HORSE::
/BASE_PRIORITY=4 /JOB_LIMIT=1 /OWNER=[SYSTEM]
/PROTECTION=(S:E,O:D,G:R,W:W)
    .
    .
    .
Batch queue BAILEY_BATCH, on BAILEY::
/BASE_PRIORITY=4 /JOB_LIMIT=1 /OWNER=[SYSTEM]
/PROTECTION=(S:E,O:D,G:R,W:W)

Generic batch queue SYS$BATCH
/GENERIC=(LION_BATCH,TIGER_BATCH,BEAR_BATCH,
HORSE_BATCH,BARNUM_BATCH,RNGLNG_BATCH,BAILEY_BATCH)
/OWNER=[SYSTEM] /PROTECTION=(S:E,O:D,G:R,W:W)
```

Sample Queue Startup Command Procedures

Example 5–17 Queue Startup Command Procedure for BARNUM

```
$ SET NOON
!
! STARTQUE.COM for BARNUM
!
! Start job queue manager
!
$ START/QUE/MANAGER SYS$COMMON:[SYSEXE]JBCSYSQUE.DAT
!
! Initialize and start local print queues
!
$ INITIALIZE/QUEUE/ON=BARNUM::LPA0:/START BARNUM_PRINT
!
! Initialize remote print queues
!
$ INITIALIZE/QUEUE/ON=BAILEY::LPA0: BAILEY_PRINT
!
! Initialize and start cluster-wide generic print queue
!
$ INITIALIZE/QUEUE/GENERIC=(BARNUM_PRINT,BAILEY_PRINT) -
  /START SYS$PRINT
!
! Initialize and start local batch queues
!
$ INITIALIZE/QUEUE/BATCH/ON=BARNUM::/START BARNUM_BATCH
!
! Initialize remote batch queues
!
$ INITIALIZE/QUEUE/BATCH/ON=LION::          LION_BATCH
$ INITIALIZE/QUEUE/BATCH/ON=TIGER::         TIGER_BATCH
$ INITIALIZE/QUEUE/BATCH/ON=BEAR::          BEAR_BATCH
$ INITIALIZE/QUEUE/BATCH/ON=HORSE::         HORSE_BATCH
$ INITIALIZE/QUEUE/BATCH/ON=RNGLNG::        RNGLNG_BATCH
$ INITIALIZE/QUEUE/BATCH/ON=BAILEY::        BAILEY_BATCH
!
! Initialize and start cluster-wide generic batch queue
!
$ INITIALIZE/QUEUE/BATCH/GENERIC=(LION_BATCH,TIGER_BATCH,BEAR_BATCH,-
  HORSE_BATCH,BARNUM_BATCH,RNGLNG_BATCH,BAILEY_BATCH)/START  SYS$BATCH
```

Common Queue Startup

Example 5–18 Common Command Procedure for Queue Startup

```
$ SET NOON
!
! STARTQUE.COM for all nodes
!
! Initialize Symbols
$ LION_START="/NOSTART"
$ TIGER_START="/NOSTART"
$ BEAR_START="/NOSTART"
$ HORSE_START="/NOSTART"
$ BARNUM_START="/NOSTART"
$ RNGLNG_START="/NOSTART"
$ BAILEY_START="/NOSTART"
!
! Set symbol for this member
!
$ NODE = F$GETSYI("NODENAME")
$ 'NODE'_START = "/START"
!
! Start job queue manager
!
$ START/QUE/MANAGER MAN_DSK:[CLUSMAN]JBCSYSQUE.DAT
!
! Initialize and start cluster print queues (these are the only two members
! that have local printers)
!
$ INIT/QUE/ON=BARNUM::LPA0:'BARNUM_START BARNUM_PRINT
$ INIT/QUE/ON=BAILEY::LPA0:'BAILEY_START BAILEY_PRINT
!
! Initialize and start cluster-wide generic print queue. All members in the
! cluster need to start this generic queue
!
$ INIT/QUE/GENERIC=(BARNUM_PRINT,BAILEY_PRINT)/START SYS$PRINT
!
! Initialize and start batch queues
$ INITIALIZE/QUEUE/BATCH/ON=BARNUM::'BARNUM_START  BARNUM_BATCH
$ INITIALIZE/QUEUE/BATCH/ON=LION::'LION_START      LION_BATCH
$ INITIALIZE/QUEUE/BATCH/ON=TIGER::'TIGER_START    TIGER_BATCH
$ INITIALIZE/QUEUE/BATCH/ON=BEAR::'BEAR_START      BEAR_BATCH
$ INITIALIZE/QUEUE/BATCH/ON=HORSE::'HORSE_START    HORSE_BATCH
$ INITIALIZE/QUEUE/BATCH/ON=RNGLNG::'RNGLNG_START  RNGLNG_BATCH
$ INITIALIZE/QUEUE/BATCH/ON=BAILEY::'BAILEY_START  BAILEY_BATCH
!
! Initialize and start cluster-wide generic batch queue
!
$ INITIALIZE/QUEUE/BATCH/GENERIC=(LION_BATCH,TIGER_BATCH,BEAR_BATCH,-
  HORSE_BATCH,BARNUM_BATCH,RNGLNG_BATCH,BAILEY_BATCH)/START  SYS$BATCH
```

Starting Up the LAT (Local Area Transport) Environment

LAT$STARTUP

SYS$MANAGER:~~LTLOAD~~.COM

- Initializes the LAT protocol on a VMS service node
- Allows service nodes (such as cluster members) to offer services to devices connected to terminal servers, which are connected to the Ethernet
- Executed after STARTNET.COM
- Is executed from SYSTARTUP_V5.COM, either directly or submitted for batch execution

By editing the command procedure ~~LTLOAD~~ LAT$STARTUP.COM, you can define how the LAT is going to work, including such complex configurations as a cluster service that balances the load among cluster members.

To create a configuration where the VMS member supports interactive terminals:

- A single command line in the procedure where LTLOAD is called is sufficient.

  LAT$STARTUP
  ```
  $ @SYS$MANAGER:LTLOAD "CIRCUS" "/ENABLE=(1,21,143)" "/DISABLE=0" ""
  ```

  — By default, ~~LTLOAD~~ LAT$STARTUP will create a service with the name given by SYS$NODE.

  — This command creates an additional service, CIRCUS (from the P1 parameter), which is accessible from terminal servers with the group codes 1, 21, 143.

  — The server manager can assign group codes to servers or specific terminals.

LAT$STARTUP.COM

Example 5–19 Edited LTLOAD.COM Procedure for an Entire Cluster

```
$ !  Copyright (c) 1987 Digital Equipment Corporation.  All rights reserved.
$ ! This command procedure starts up the LAT protocol
$ ! and configures applications devices for remote printer use.
$
$ RUN SYS$SYSTEM:SYSGEN
CONNECT LTA0/NOADAPTER
$
$! Invoke LATCP
$
$LCP := $LATCP
!
! The following commands will set up LAT service with the default name
! SYS$NODE and default ident SYS$ANNOUNCE.  The LAT service name will
! default to the node name SYS$NODE unless you specify the name as
! the first parameter in the command line.  Additional node characteristics
! such as group codes can also be supplied as parameters.
!
$LCP SET NODE /IDENT  'P2' 'P3' 'P4' /NOLOG
$LCP CREATE SERVICE /ID
$IF P1 .NES. "" THEN $LCP CREATE SERVICE 'P1' /IDENT /NOLOG
!
! *** This line, using the P1 parameter, creates service "CIRCUS"
!      for the Circus cluster ***
! *** This service name will be in addition to the default SYS$NODE
!      service set up in the previous line.
$!
$ RUN SYS$SYSTEM:LATCP
!
! Set up the applications devices that will support remote printer
! access.
!
! Create the devices.
!
CREATE PORT LTA1: /NOLOG
CREATE PORT LTA2: /NOLOG
!
! Maps applications port(s) to a specific port(s) on the terminal
! server
!
SET PORT LTA1: /APPLICATION /NODE=SERVER_1 /PORT=LN03
SET PORT LTA2: /APPLICATION /NODE=SERVER_2 /PORT=LN03R
!
!
! Start LAT Service
!
START NODE
EXIT
! *** Because this procedure is usually executed after the queues have been
! started, we will set up the terminal and queue here, and start it for
! the LAT queues. ***
$ SET TERMINAL LTA1: /PERM /DEVICE=LN03 /WIDTH=255 /PAGE=60 /LOWERCASE /NOBROAD
$ SET TERMINAL LTA2: /PERM /DEVICE=LN03R/WIDTH=255 /PAGE=60 /LOWERCASE /NOBROAD
$ SET DEVICE LTA1: /SPOOLED=(LN03$PRINT, SYS$SYSDEVICE)
$ SET DEVICE LTA2: /SPOOLED=(LN03R$PRINT, SYS$SYSDEVICE)
$ INIT/QUEUE/START/PROCESSOR=LATSYM/ON=BARNUM::LTA1:     LN03_PRINT
$ INIT/QUEUE/START/PROCESSOR=LATSYM/ON=BARNUM::LTA2:     LN03R_PRINT
```

CONFIGURING DISK AND TAPE VOLUMES

Tape Drive Access

To define accessibility of tape devices:

- Connect the device directly to the node for local access.
- Connect the device through the HSC unit for cluster-wide access.

Disk Access

To mount disk volumes in a common environment VAXcluster system:

- Make local devices available cluster-wide through the MSCP server.
- Mount HSC based disks on all CI members.
- Mount DSSI based ISEs on all DSSI members.
- Mount MSCP served HSC, DSSI, and local disks on any or all nodes.
  - The command MOUNT/CLUSTER mounts a disk volume on all nodes currently in the VAXcluster system.
  - A node that joins the cluster later must explicitly mount the volume.
- To ensure that all volumes remain mounted cluster-wide in startup command procedure:
  - Mount all local MSCP served disks with the MOUNT/CLUSTER command.
  - Mount all remote disks with the MOUNT or MOUNT/CLUSTER command.

Making a Dual-Ported Disk Available Cluster-Wide

- Be sure that the MSCP server is loaded (MSCP_LOAD = 1) and enabled for serving local disks (MSCP_SERVE_ALL = 1 or 2)
- In your startup command procedure:
  - — If the disk is a MASSBUS disk, enable it for dual-porting on each node

    ```
    $ SET DEVICE/DUAL_PORT $1$DRA1:
    ```

  - — Mount the disk

    ```
    $ MOUNT/CLUSTER $1$DRA1: COMMONPACK
    ```

Sample Command Procedures to Mount Disks

Example 5-20 Command Procedure to Mount Disks for HORSE

```
$ SET NOON
!
! MOUNT.COM -- Command procedure to mount disks for HORSE
!
! Mount local disks
!
$ MOUNT /CLUSTER $2$DUA0 THREE
$ MOUNT /CLUSTER $2$DUA1 RING
!
! Mount remote disks
!
$ MOUNT/SYSTEM/NOASSIST $1$DUA1 TIGHTROPE
$ MOUNT/SYSTEM/NOASSIST $1$DUA2 FLYING
$ MOUNT/SYSTEM/NOASSIST $1$DUA3 TRAPEZE
!
$ EXIT
```

Example 5-21 Command Procedure to Mount Disks for BAILEY

```
$ SET NOON
!
! MOUNT.COM -- Command procedure to mount disks for BAILEY
!
! Mount local disks
!
$ MOUNT /CLUSTER $1$DUA3 TRAPEZE
!
! Mount remote disks
!
$ MOUNT /SYSTEM $1$DUA1 TIGHTROPE
$ MOUNT /SYSTEM $1$DUA2 FLYING
$ MOUNT /SYSTEM $2$DUA0 THREE
$ MOUNT /SYSTEM $2$DUA1 RING
!
$ EXIT
```

Rebuilding Incorrectly Dismounted System Disks

Free space and storage allocation inconsistencies caused by incorrect dismounting of the system disk can often be fixed by rebuilding the disk.

Rebuild considerations include:

- While the rebuild is in progress, the disk is unavailable.
- If the system disk is rebuilt at boot time, and the system disk is also being served by the MSCP server, there is the potential for the cluster to hang.
- To prevent this situation on any system that serves a system disk, or boots from a served system disk, you should do the following:
  - Edit MODPARAMS.DAT and execute AUTOGEN to set the SYSGEN parameter ACP_REBLDSYSD to 0
  - Have the system disk rebuilt at a more convenient time (such as in a batch job at an off-hour)

    ```
    $ SET VOLUME/REBUILD SYS$SYSDEVICE
    ```
- There is no risk to data integrity by not rebuilding the system disk right away. The worst consequence is that some free space will not be available until the disk is rebuilt.
- CLUSTER_CONFIG sets ACP_REBLDSYSD to 0 for satellites only. The system manager must set it for boot servers.

Using MSCPMOUNT.COM to Wait for Check for Mountable Disks

I/O to a remote disk when the remote system serving the disk shuts down or becomes unavailable results in:

- The disk going into mount verification
- I/O resuming if the system reboots within the mount verification limit
- A need to remount the disk if the system fails to reboot within that time

The command procedure SYS$EXAMPLES:MSCPMOUNT.COM can be used to automatically remount disks in this state:

- Checks every 15 minutes for disks that have become unavailable
- Mounts any available disks that are not already mounted
- Dismounts and remounts any disks for which mount verification has timed out

Edit MSCPMOUNT.COM to conform to your site's configuration:

- Remove names of example disks
- Add names of disks in your environment

```
$ RECURSIVE_CALLBACK MOUNT $1$DUA0 BARNUM_SYS
$ RECURSIVE_CALLBACK MOUNT $1$DUA1 TIGHTROPE
$ RECURSIVE_CALLBACK MOUNT $1$DUA2 FLYING
$ RECURSIVE_CALLBACK MOUNT $1$DUA3 TRAPEZE
```

- Place the edited version in SYS$SPECIFIC:[SYSMGR] of each node
- Place the following command in the startup command procedure

```
$ @SYS$SPECIFIC:[SYSMGR]MSCPMOUNT.COM
```

Proper Dismount of Disks on Shutdown

Building a VAXcluster system should include setting up SYS$MANAGER:SYSHUTDWN.COM.

When you shut down a system:

- DISMOUNT/CLUSTER each MSCP served disk on the system that is not dual-ported.
  - Disks that are not dismounted will undergo mount verification on other systems that have them mounted.
  - All processes with outstanding I/O to these disks will hang.
  - If mount verification times out, the other systems must dismount and remount the disks.
- Dismount disks in the site-specific shutdown procedure SYS$MANAGER:SYSHUTDWN.COM
  - For example, on HORSE:

    ```
    $ DISMOUNT/CLUSTER/ABORT $2$DUA0:
    ```

  - /ABORT forces a disk to be dismounted even if it has open user files. This will not work if the open files are page or swap files, or if the disk is a system disk.

SUMMARY

- To build a VAXcluster system:
  - Install VMS operating system software and configure DECnet on one system disk.
  - Use CLUSTER_CONFIG to add roots to the system disk and create additional system disks.
  - Create bootstrap command procedures to allow nodes to boot into the cluster automatically.
  - Create coordinated startup command procedures to mount disks and start queues.
- When merging existing systems to form a cluster:
  - Coordinate UAF, mail, proxy, and rights databases.

APPENDIX A — ADDITIONAL HARDWARE COMMANDS

HSC SETSHO Commands

Table 5–6 SETSHO Set Commands

Command	Parameters	Reboot
ALLOCATE DISK	Allocation-class	Yes
ALLOCATE TAPE	Allocation-class	Yes
AUTOMATIC DIAGNOSIS	Sense	Yes
CI	Node-number	No
DATE	Date-and-Time	No
DUMP	Sense Device	No
ERROR	Error-level	No
HOST ENABLE/DISABLE	Node-number	Yes
ID	System-id	Yes
LOAD	Load-module	No
MAX_FORMATTER	Formatter-count	Yes
MAX_TAPES	Tape-count	Yes
MEMORY ENABLE ENABLE/ALL DISABLE	Memory-address	Yes
NAME	System-name	Yes
ODT	Sense mode	Yes
OUTBAND	Error-level	No
PERIODIC_DIAGNOSTICS	Interval	Yes
RESTART		No
SCOPE		No
SCT CLEAR		Yes
SECTOR_SIZE	Sector_size	Yes
SERVER DISK/DRIVE_TIMEOUT	Speed of Failover	No
UNIT_ID	unit-id	No

MicroVAX 3300 and MicroVAX 3400 Console Commands

Table 5–7 MicroVAX 3300 and MicroVAX 3400 Console Command Summary

Command	Qualifiers	Argument	Other
BOOT	/R5:{bitmap}/{bitmap}	[{device_ string}]	
CONTINUE			
DEPOSIT	/B/W/L/Q/G/I/V/P/M/U/N:{count}	{address}	{data}[{data}]
	STEP:{size}/WRONG		
EXAMINE	/B/W/L/Q/G/I/V/P/M/U/N:{count}	{address}	
	STEP:{size}/WRONG/INSTRUCTION		
FIND			
HALT			
HELP			
INITIALIZE			
MOVE	/B/W/L/Q/V/P/U/N:{count}	{src_address}	{dest_ address}
NEXT		[{count}]	
REPEAT		{command}	
SEARCH	/B/W/L/Q/V/P/U/N:{count}	{start_ address}	{pattern}
	/STEP:{size}/WRONG/NOT		
SET BFLAG		{bitmap}	
SET BOOT		{device_strng}	
SET HOST	DUP{DSSI!/UQSSP}{/DISK!TAPE!csr}	{node} !	[{task}]
	MAINTENANCE/UQSSP{ /SERVICE!csr}	{controller_ number}	

Table 5–7 Console Command Summary (Cont.)

Command	Qualifiers	Argument	Other
SET LANGUAGE		{language type}	
SHOW BFLAG			
SHOW BOOT			
SHOW DSSI			
SHOW ETHERNET			
SHOW LANGUAGE			
SHOW MEMORY	FULL		
SHOW QBUS			
SHOW RLV12			
SHOW VERSION			
START		{address}	
TEST		{address}	
UNJAM			
X		{address}	{count}

SET and SHOW User Parameters for ISEs

Table 5–8 ISE SET and SHOW User Parameters

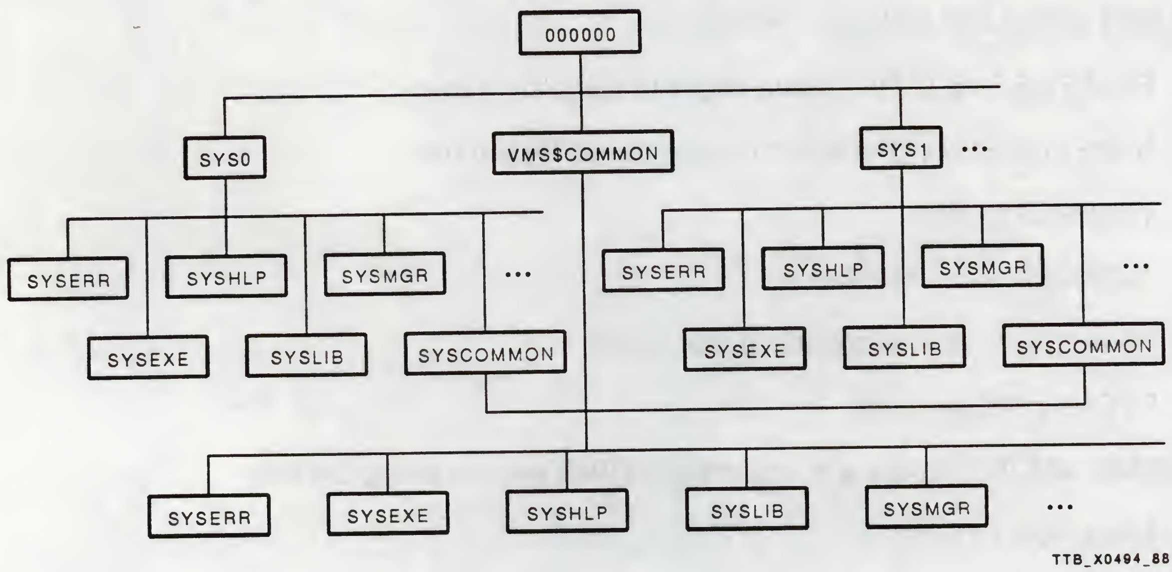
Parameter	Class	Definition
VOLSERNO	DRIVE	Volume serial number as a quadword
ALLCLASS	MSCP	Controller allocation class. Must match host.
UNITNUM	MSCP	MSCP unit number
FORCEUNI	MSCP	Determines whether MSCP unit number or DSSI node ID is used: 0 = MSCP, 1 = DSSI
FIVEDIME	MSCP	Credit connections. 0 = seven connections with seven credits each, 1 = five connections with ten credits each.
CNT_TMO	MSCP	MSCP controller timeout value
ADD_CR	DUP	Determines if DUP appends to a LINEFEED character after each message.

APPENDIX B — SYSTEM DISK DIRECTORY STRUCTURE

- [SYSn.*] contains all node-specific files:
  - PAGEFILE.SYS (if the system pages to the system disk)
  - SWAPFILE.SYS (if the system swaps to the system disk)
  - VAXVMSSYS.PAR
  - AUTOGEN parameter files
  - Node-specific startup command procedures
  - DECnet database files
- [VMS$COMMON.*] contains all shared files (VMS and optional products):
  - Executable images
  - Libraries
  - Common startup command procedures
- [SYS0.*] and [VMS$COMMON.*] are created by initial installation of the VMS operating system
- [SYSn.*] is created by SYS$MANAGER:CLUSTER_CONFIG.COM
- Each SYSn root contains [SYSn.SYSCOMMON]
- [SYSn.SYSCOMMON] is an alias for [VMS$COMMON]

Common System Disk Structure

Figure 5–6 Common System Disk Structure



Logical Names and the Common System Disk Directory Structure

- The logical name SYS$SPECIFIC points to the node-specific root from which the system is booted

  ```
  SYS$SPECIFIC = SYS$SYSDEVICE:[SYSn.]
  ```

  If we were booting from 1DUA0, root [SYS1]:

  ```
  SYS$SPECIFIC = $1$DUA0:[SYS1.]
  ```

- The logical name SYS$COMMON points to the common directory tree

  ```
  SYS$COMMON = SYS$SYSDEVICE:[SYSn.SYSCOMMON.]
  ```

  Using the example above:

  ```
  SYS$COMMON = $1$DUA0:[SYS1.SYSCOMMON.]
  ```

- SYS$SYSROOT is a VAX RMS search list

  ```
  SYS$SYSROOT = SYS$SYSDEVICE:[SYSn.]
              = SYS$COMMON:
  ```

  Using the example above:

  ```
  SYS$SYSROOT = $1$DUA0:[SYS1.]
              = SYS$COMMON:
  ```

- Some logical names point to two directories, for example:

  ```
  SYS$SYSTEM = SYS$SYSROOT:[SYSEXE]
  ```

  In the example above, SYS$SYSTEM translates to both:

  ```
  $1$DUA0:[SYS1.SYSEXE], $1$DUA0:[SYS1.SYSCOMMON.SYSEXE]
  ```

  It is important to keep this search order in mind when manipulating system files. When opening a file, the VMS operating system searches the node-specific root first.

- To refer to a single directory, use SYS$SPECIFIC or SYS$COMMON rather than SYS$SYSROOT

- To edit your node's own site-specific startup file:

```
$ EDIT SYS$SPECIFIC:[SYSMGR]SYSTARTUP_V5.COM
```

- To edit another node's own site-specific startup file from your node:

```
$ EDIT $1$DUA0:[SYS2.SYSMGR]SYSTARTUP_V5.COM
```

(assuming 1DUA0 and [SYS2] are the disk and root from which the other system boots)

Example 5–22 Some Logical Names Related to the Directory Structure

```
$ SHOW LOGICAL SYS$*
   .
   .
   .
   "SYS$COMMON" = "$1$DUA0:[SYS0.SYSCOMMON.]"
   "SYS$DISK" = "$1$DUA0:"
   "SYS$ERRORLOG" = "SYS$SYSROOT:[SYSERR]"
   "SYS$EXAMPLES" = "SYS$SYSROOT:[SYSHLP.EXAMPLES]"
   "SYS$HELP" = "SYS$SYSROOT:[SYSHLP]"
   "SYS$INSTRUCTION" = "SYS$SYSROOT:[SYSCBI]"
   "SYS$LIBRARY" = "SYS$SYSROOT:[SYSLIB]"
   "SYS$LOADABLE_IMAGES" = "SYS$SYSROOT:[SYS$LDR]"
   "SYS$MAINTENANCE" = "SYS$SYSROOT:[SYSMAINT]"
   "SYS$MANAGER" = "SYS$SYSROOT:[SYSMGR]"
   "SYS$MESSAGE" = "SYS$SYSROOT:[SYSMSG]"
   "SYS$SHARE" = "SYS$SYSROOT:[SYSLIB]"
   "SYS$SPECIFIC" = "$1$DUA0:[SYS0.]"
   "SYS$STARTUP" = "SYS$SYSROOT:[SYS$STARTUP]"
         = "SYS$MANAGER"
   "SYS$SYLOGIN" = "SYS$MANAGER:SYLOGIN.COM"
   "SYS$SYSDEVICE" = "$1$DUA0:"
   "SYS$SYSDISK" = "SYS$SYSROOT:"
   "SYS$SYSROOT" = "$1$DUA0:[SYS0.]"
         = "SYS$COMMON:"
   "SYS$SYSTEM" = "SYS$SYSROOT:[SYSEXE]"
   "SYS$TEST" = "SYS$SYSROOT:[SYSTEST]"
   "SYS$TOPSYS" = "SYS0"
   "SYS$UPDATE" = "SYS$SYSROOT:[SYSUPD]"
   "SYS$WELCOME" = "@SYS$MANAGER:WELCOME.TXT"
```

SYS$SPECIFIC Directory Structure

Example 5-23 SYS$SPECIFIC Directory Structure

```
$ DIRECTORY SYS$SYSDEVICE:[000000]

Directory SYS$SYSDEVICE:[000000]

000000.DIR;1        BACKUP.SYS;1        BADBLK.SYS;1        BADLOG.SYS;1
BITMAP.SYS;1        CONTIN.SYS;1        CORIMG.SYS;1        COURSE.DIR;1
INDEXF.SYS;1        SYS0.DIR;1          SYS1.DIR;1          SYS10.DIR;1
SYS11.DIR;1         SYSE.DIR;1          SYSEXE.DIR;1        SYSLOST.DIR;1
SYSMAINT.DIR;1      VMS$COMMON.DIR;1    VOLSET.SYS;1

Total of 19 files.
$ DIRECTORY SYS$SYSDEVICE:[SYS0]

Directory SYS$SYSDEVICE:[SYS0]

DECNET.DIR;1        MOM$SYSTEM.DIR;1    SYS$LDR.DIR;1       SYS$STARTUP.DIR;1
SYSCBI.DIR;1        SYSCOMMON.DIR;1     SYSERR.DIR;1        SYSEXE.DIR;1
SYSHLP.DIR;1        SYSLIB.DIR;1        SYSMAINT.DIR;1      SYSMGR.DIR;1
SYSMSG.DIR;1        SYSTEST.DIR;1       SYSUPD.DIR;1

Total of 15 files.
$ DIRECTORY SYS$SYSDEVICE:[SYS0.SYSMGR]

Directory SYS$SYSDEVICE:[SYS0.SYSMGR]

ACCOUNTNG.DAT;39    LTLOAD.COM          MOUNT.COM;2
NETCONFIG.COM;1     OPERATOR.LOG;108    STARTNET.COM;43
STARTQUE.COM;22     SYCONFIG.COM;5      SYSHUTDOWN.COM;2     . . .
SYSTARTUP_V5.COM;330                    TERMINALS.COM;3
VMSIMAGES.DAT;6

Total of 12 files.
$ DIRECTORY SYS$SPECIFIC:[SYSMGR]

Directory SYS$SPECIFIC:[SYSMGR]

ACCOUNTNG.DAT;39    LTLOAD.COM          MOUNT.COM;2
NETCONFIG.COM;1     OPERATOR.LOG;108    STARTNET.COM;43
STARTQUE.COM;22     SYCONFIG.COM;5      SYSHUTDOWN.COM;2     . . .
SYSTARTUP_V5.COM;330                    TERMINALS.COM;3
VMSIMAGES.DAT;6

Total of 12 files.
```

SYS$COMMON Directory Structure

Example 5–24 SYS$COMMON Directory Structure

```
$ DIRECTORY SYS$SYSDEVICE:[SYS0.SYSCOMMON]/FILE

Directory SYS$SYSDEVICE:[SYS0.SYSCOMMON]

MOM$SYSTEM.DIR;1      (3262,2,0)
SYS$LDR.DIR;1         (3263,2,0)
SYS$STARTUP.DIR;1     (3264,2,0)
SYSCBI.DIR;1          (3260,2,0)
SYSERR.DIR;1          (3259,2,0)
SYSEXE.DIR;1          (1472,5,0)
SYSFONT.DIR;1         (2357,7,0)
SYSHLP.DIR;1          (3257,2,0)
SYSLIB.DIR;1          (2458,11,0)
SYSMAINT.DIR;1        (3256,2,0)
SYSMGR.DIR;1          (3255,3,0)
SYSMSG.DIR;1          (3254,7,0)
SYSTEST.DIR;1         (3258,2,0)
SYSUPD.DIR;1          (3261,2,0)

Total of 14 files.
$ DIRECTORY SYS$SYSDEVICE:[VMS$COMMON]/FILE

Directory SYS$SYSDEVICE:[VMS$COMMON]

MOM$SYSTEM.DIR;1      (3262,2,0)
SYS$LDR.DIR;1         (3263,2,0)
SYS$STARTUP.DIR;1     (3264,2,0)
SYSCBI.DIR;1          (3260,2,0)
SYSERR.DIR;1          (3259,2,0)
SYSEXE.DIR;1          (1472,5,0)
SYSFONT.DIR;1         (2357,7,0)
SYSHLP.DIR;1          (3257,2,0)
SYSLIB.DIR;1          (2458,11,0)
SYSMAINT.DIR;1        (3256,2,0)
SYSMGR.DIR;1          (3255,3,0)
SYSMSG.DIR;1          (3254,7,0)
SYSTEST.DIR;1         (3258,2,0)
SYSUPD.DIR;1          (3261,2,0)

Total of 14 files.
```

Managing VAXcluster Operations

INTRODUCTION

This module covers only those areas where operating or reconfiguring a VAXcluster system differs from operating or reconfiguring a single VAX system.

OBJECTIVES

To manage a VAXcluster system, the system manager should be familiar with:

- The SYSMAN utility
- Installation and update procedures of layered products in a cluster
- Managing licenses in a cluster
- Managing operator communications
- Reconfiguring disks, HSC subsystems, and VMS members
- Performing backups of VAXcluster disks
- VMS update procedures
- The procedure to restore cluster quorum after node failure
- Execution of conditional shutdown operations
- Security functions in a local area or mixed-interconnect VAXcluster system

RESOURCES

- *VMS DCL Dictionary*
- *Guide to Maintaining a VMS System*
- *VMS VAXcluster Manual*
- *VMS System Generation Utility Manual*
- *VMS Volume Shadowing Manual*
- *VAX Volume Shadowing Manual*
- *Introduction to VMS System Services*
- *VMS System Services Reference Manual*
- *VMS License Management Utility Manual*
- VMS Installation and Operations manual for your processor

SYSTEM MANAGEMENT (SYSMAN) UTILITY

The System Management utility provides a system manager a comprehensive tool to perform many typical system management tasks.

SYSMAN allows the system manager to define the environment in which these tasks are performeo to be:

- The node the user is logged in to
- Some set of nodes in the user's own cluster
- All nodes in a cluster
- Some other single node in the cluster
- Some other node in another cluster by means of the DECnet network
- Some set of nodes in another cluster by means of the DECnet network
- A standalone node by means of the DECnet network

SYSMAN allows DCL, DISKQUOTA, and SYSGEN commands to be issued by any cluster member.

Implementation of SYSMAN commands:

- Execute in the context of the SMI server process on each member
- Use SCS for cluster communication
- Use the DECnet network for communication outside the cluster

To invoke SYSMAN:

- $ RUN SYS$SYSTEM:SYSMAN
- SYSMAN can be defined as a foreign command

```
$ SYSMAN == "$SYS$SYSTEM:SYSMAN"
```

SYSMAN Command Summary

Table 6–1 summarizes the various SYSMAN commands.

Table 6–1 SYSMAN Commands and Subcommands

Command	Qualifier	Function
SET ENVIRONMENT	/CLUSTER	All nodes in the cluster are target.
	/NODE=nodename	Specifies a remote node as target.
	/CLUSTER /NODE=node	Specifies a remote cluster.
	/USERNAME=username	Unless specified, attempts to log in to a remote node with the current user name.
SHOW ENVIRONMENT		Displays the target node(s) or cluster where SYSMAN executes commands.
SET PROFILE		Temporarily changes your privileges, default device, and/or default directory in the current environment.
	/DEFAULT= device:[directory]	Modifies default device and/or directory. (Optional qualifier.)
	/PRIVILEGES= (priv1,priv2...)	Modifies process privileges. (Optional qualifier.)
SHOW PROFILE		Displays the current privileges and the default device and directory being used in the current environment.
SET TIMEOUT time		Sets the time SYSMAN will wait for a node to respond to a command issued from the current process time = delta time
PARAMETERS x		Displays or modifies system (SYSGEN) parameters in the current environment. x = SYSGEN parameter command.

Table 6-1 SYSMAN Commands (Cont.)

Command	Qualifier	Function
CONFIGURATION x		Displays or modifies certain cluster parameters in the current environment.
DISKQUOTA x		Displays or modifies disk quotas. x = DISKQUOTA command.
	/DEVICE=device	Specifies disk on which to perform DISKQUOTA function. (Optional qualifier.)
DO x		x = a DCL command. Executes a DCL command or command procedure on all nodes in an environment. Each DO command creates a new subprocess.
	/OUTPUT=filespec	If you omit the file specification, the output is written to SYSMAN.LIS in the current device and directory. (Optional qualifier.)
HELP		Displays on-line HELP
EXIT		Exits SYSMAN

SYSMAN Communication

Within a cluster:

- SMISERVER is run as a detached process.
- The SMISERVER process or subprocess is responsible for executing SYSMAN commands on remote nodes.
- The DECnet network is used for communication to a remote node.
  - SMISERVER is declared as a network object.

Between clusters:

- DECnet communication is used to the first node in the remote cluster.

On any node that is part of a cluster (the SYSGEN parameter VAXCLUSTER has a value of 1 or 2), SMISERVER is normally started by the system startup procedure SYS$SYSTEM:STARTUP.COM.

SYSMAN Commands

The SET ENVIRONMENT Command

- Requires OPER or SETPRV privilege on all nodes in the target environment
- Allows you to change the environment to any combination of the following:
  - Any other node in the cluster
  - The entire cluster
  - Any node or entire cluster available through the DECnet network

SET ENVIRONMENT Qualifiers

- /CLUSTER
  - Default management environment is the local cluster
  - Specify a remote cluster by naming one cluster member with the /NODE qualifier
- /NODE[=nodename]
  - Node name can be a system name or a cluster alias
- /USERNAME=username
  - Prompts with NOECHO for remote password
  - Does not request password inside cluster
  - Must be an authorized user on the remote system

Some DCL commands operate cluster-wide by design

- MOUNT/CLUSTER and SET CLUSTER/EXPECTED_VOTES are examples of this type of command.
- Do not use SYSMAN to execute these commands.

Example 6–1 SYSMAN SET/SHOW ENVIRONMENT

```
❶ SYSMAN> SHOW ENVIRONMENT
%SYSMAN-I-ENV, current command environment:
 Local node only

❷ SYSMAN> SET ENVIRONMENT/NODE=BARNUM
%SYSMAN-I-ENV, current command environment:
 Individual nodes: BARNUM
 Username PRIDE       will be used on nonlocal nodes

❸ SYSMAN> SHOW ENVIRONMENT
%SYSMAN-I-ENV, current command environment:
 Individual nodes: BARNUM
 Username PRIDE       will be used on nonlocal nodes

❹ SYSMAN> SET ENVIRONMENT/CLUSTER
%SYSMAN-I-ENV, current command environment:
 Clusterwide on local cluster
 Username PRIDE       will be used on nonlocal nodes

SYSMAN> SHOW ENVIRONMENT
%SYSMAN-I-ENV, current command environment:
 Clusterwide on local cluster
 Username PRIDE       will be used on nonlocal nodes
```

Notes on Example 6–1:

❶ Show the current environment.

❷ Set the environment to a remote node.

- The current user name is displayed.
- No password is requested because the user has an account on the specified node.

❸ Show the new environment.

❹ The target environment consists of the local cluster.

The SET PROFILE Command

- Enables you to modify two attributes of the SMISERVER subprocess that execute commands for you in the target environment:
  - — Current privileges (any enhanced privileges must be authorized)
  - — Default device and/or directory
- This profile is in effect until you change it, reset the environment, or exit from SYSMAN.
- In clusters where SYSUAF and RIGHTSLIST are common, instead of requiring a lookup on each node, SYSMAN grants the privileges that are in effect on the local node if the following conditions are met:
  - — The node where the commands are being entered is part of the same cluster as the node where the commands are being executed.
  - — Device name and file ID for SYSUAF match on both nodes.
  - — Device name and file ID for RIGHTSLIST match on both nodes.
  - — If not met, SYSMAN grants only the privileges in the remote system's SYSUAF and RIGHTSLIST files.

Example 6-2 SYSMAN SET/SHOW PROFILE (VAXcluster System)

```
❶ SYSMAN> SHOW PROFILE
%SYSMAN-I-DEFDIR, default directory on node BAILEY -- BAILEY$DJA0:[PRIDE]
%SYSMAN-I-DEFPRIV, process privileges on node BAILEY --
 SETPRV
 TMPMBX
 OPER
 NETMBX
 SYSPRV
❷ SYSMAN> SET ENVIRONMENT/CLUSTER
%SYSMAN-I-ENV, current command environment:
 Clusterwide on local cluster
 Username PRIDE will be used on nonlocal nodes
❸ SYSMAN> SHOW PROFILE
%SYSMAN-I-DEFDIR, default directory on node BAILEY -- BAILEY$DJA0:[PRIDE]
%SYSMAN-I-DEFPRIV, process privileges on node BAILEY --
 SETPRV
 TMPMBX
 OPER
 NETMBX
 SYSPRV
%SYSMAN-I-DEFDIR, default directory on node BARNUM -- BAILEY$DJA0:[PRIDE]
%SYSMAN-I-DEFPRIV, process privileges on node BARNUM --
 SETPRV
 TMPMBX
 NETMBX
❹ SYSMAN> SET PROFILE/PRIVILEGE=CMKRNL
SYSMAN> SHOW PROFILE
%SYSMAN-I-DEFDIR, default directory on node BAILEY -- BAILEY$DJA0:[PRIDE]
%SYSMAN-I-DEFPRIV, process privileges on node BAILEY --
 CMKRNL
 SETPRV
 TMPMBX
 OPER
 NETMBX
 SYSPRV
❺ %SYSMAN-I-DEFDIR, default directory on node BARNUM -- BAILEY$DJA0:[PRIDE]
%SYSMAN-I-DEFPRIV, process privileges on node BARNUM --
 CMKRNL
 SETPRV
 TMPMBX
 NETMBX
```

Notes on Example 6-2:

❶ Profile before any changes have been made. Note that the user has five privileges.

❷ The target environment is the cluster containing BARNUM and BAILEY. No password prompt

❸ Profile information is displayed for all nodes in the specified cluster.

❹ A new privilege, CMKRNL, is added to the profile.

❺ The new privilege takes effect cluster-wide.

The PARAMETERS Command

The PARAMETERS command lets you perform the following actions on system parameters and parameter files:

- Inspect
- Set
- Modify

The format for the SYSMAN PARAMETERS command is:

```
SYSMAN> PARAMETERS subcommand
```

Table 6–2 SYSMAN PARAMETERS Subcommands

Subcommand	Function
SET	Modifies the value of a system parameter in the work area.
SHOW	Displays the values of system parameters in the work area, plus the default, minimum, and maximum values of the parameters and their units of measure.
USE	Initializes the current work area with system parameter values.
WRITE	Writes the system parameter values to a parameter file, to the current system parameter file, or to the active system in memory.

Example 6–3 SYSMAN PARAMETERS Command

```
❶ SYSMAN> SET ENVIRONMENT/NODE=BARNUM
  %SYSMAN-I-ENV, current command environment:
    Individual nodes: BARNUM
    Username PRIDE       will be used on nonlocal nodes

❷ SYSMAN> SET PROFILE/PRIVILEGE=CMEXEC

❸ SYSMAN> PARAMETERS SHOW /ACP
❹ %SYSMAN-I-USEACTNOD, a USE ACTIVE has been defaulted on node BARNUM
  Node BARNUM:   Parameters in use: ACTIVE
  Parameter Name              Current    Default    Minimum    Maximum Unit   Dynamic
  --------------              -------    -------    -------    ------- ----   -------
  ACP_MULTIPLE                      0          0          0          1 Boolean      D
  ACP_SHARE                         1          1          0          1 Boolean
  ACP_MAPCACHE                      8          8          1         -1 Pages        D
  ACP_HDRCACHE                     36        128          3         -1 Pages        D
  ACP_DIRCACHE                     36         80          2         -1 Pages        D
  ACP_DINDXCACHE                    9         25          2         -1 Pages        D
  ACP_WORKSET                       0          0          0         -1 Pages        D
  ACP_FIDCACHE                     64         64          0         -1 File-Ids     D
  ACP_EXTCACHE                     64         64          0         -1 Extents      D
  ACP_EXTLIMIT                    100        100          0       1000 Percent/10   D
  ACP_QUOCACHE                     21         64          0         -1 Users        D
  ACP_SYSACC                        4          8          0         -1 Directories  D
  ACP_MAXREAD                      32         32          1         64 Blocks       D
  ACP_WINDOW                        7          7          1         -1 Pointers     D
  ACP_WRITEBACK                     1          1          0          1 Boolean      D
  ACP_DATACHECK                     2          2          0          3 Bit-mask     D
  ACP_BASEPRIO                      8          8          4         31 Priority     D
  ACP_SWAPFLGS                     14         15          0         15 Bit-mask     D
  ACP_XQP_RES                       1          1          0          1 Boolean
  ACP_REBLDSYSD                     1          1          0          1 Boolean
```

Notes on Example 6–3:

❶ Establish the target node.

❷ Set the privilege of the profile to enable the next command.

❸ Display the value of a group of parameters.

- You may display a specific parameter.

❹ Note that USE ACTIVE is assumed.

The CONFIGURATION Command

The format for the SYSMAN CONFIGURATION command is:

```
SYSMAN> CONFIGURATION subcommand
```

Table 6–3 SYSMAN CONFIGURATION Subcommands

Subcommand	Function
SET CLUSTER_AUTHORIZATION	Modifies security data in a mixed interconnect or local area VAXcluster system. Requires SYSPRV privilege.
SET TIME	Modifies the current system time. Requires LOG_IO privilege. In a cluster environment, also requires SYSLCK privilege.
SHOW CLUSTER_AUTHORIZATION	Displays the group number and Ethernet multicast address of a mixed interconnect or local area VAXcluster system. Requires SYSPRV privilege.
SHOW TIME	Displays the current date and system time to hundredths of a second.

The SET CLUSTER_AUTHORIZATION Subcommand

Security-related information is recorded in SYS$SYSTEM:CLUSTER_AUTHORIZE.DAT:

- Group number:
  - Uniquely identifies each local area cluster configuration on a single Ethernet
  - Must be in the range from 1 to 4095 or 61440 to 65535
- Password:
  - Prevents an intruder who discovers the group number from joining the cluster
  - Consists of 1 to 31 characters: letters, numbers, dollar sign, underscore

NOTE

If you change either the group number or the password, you must reboot the entire cluster.

The SET TIME Subcommand

- Because of communication and processing delays it is not possible to synchronize clocks exactly.
- Variation is typically less than a few hundredths of a second.
- If SYSMAN cannot set the time to within one half second of the specified time, you receive a warning message naming the node(s) that failed to respond quickly enough.

Example 6–4 SYSMAN CONFIGURATION Commands

```
❶ SYSMAN> SET ENVIRONMENT/CLUSTER
❷ SYSMAN> CONFIGURATION SHOW TIME
  System time on node BAILEY: 17-FEB-1989 12:55:37.47
  System time on node BARNUM: 17-FEB-1989 12:55:37.24

❸ SYSMAN> SET PROFILE/PRIVILEGE=SYSPRV
❹ SYSMAN> CONFIGURATION SHOW CLUSTER_AUTHORIZATION
  Node BAILEY:   Cluster group number: 65084

❺ SYSMAN> CONFIGURATION SET CLUSTER_AUTHORIZATION/PASSWORD=HOTDOG
  %SYSMAN-I-CAFOLDGROUP, existing group will not be changed
  %SYSMAN-I-CAFREBOOT, cluster authorization file updated
  The entire cluster should be rebooted
```

Notes on Example 6–4:

❶ Set the environment to be the cluster containing the node BARNUM.

❷ Display system time on all nodes in the target environment.

❸ The SHOW CLUSTER_AUTHORIZATION command requires SYSPRV privilege.

❹ The group number of node BAILEY is displayed.

- Because the group number and password on other nodes in the cluster are identical, no further information is displayed.

❺ The cluster password is modified.

- To change the group number, you must add the /GROUP_NUMBER=n qualifier to the command line.
- Note that SYSPRV was previously added to the profile.

The DISKQUOTA Command

- The format for the SYSMAN DISQUOTA comand is:

  ```
  SYSMAN> DISKQUOTA subcommand
  ```

- Performs the specified operation on the device named by the /DEVICE qualifier. If /DEVICE is not specified, your current default disk is used.

Example 6–5 SYSMAN DISKQUOTA Command

```
❶ $ SET PROCESS/PRIVILEGE=OPER
  $ RUN SYS$SYSTEM:SYSMAN

❷ SYSMAN> SET ENVIRONMENT/NODE=BARNUM
  %SYSMAN-I-ENV, current command environment:
   Individual nodes: BARNUM
   Username PRIDE       will be used on nonlocal nodes

❸ SYSMAN> SET PROFILE/PRIVILEGE=SYSPRV

❹ SYSMAN> DISKQUOTA SHOW [11,200]
  %SYSMAN-I-NODERR, error returned from node BARNUM
  %SYSTEM-F-QFNOTACT, disk quotas not enabled on this volume

❺ SYSMAN> DISKQUOTA ENABLE
  %SYSMAN-I-NODERR, error returned from node BARNUM
  %SYSTEM-W-NOSUCHFILE, no such file

❻ SYSMAN> DISKQUOTA CREATE

❼ SYSMAN> DISKQUOTA ADD [11,200]/PERM=10000/OVER=500

❽ SYSMAN> DISKQUOTA SHOW [11,200]
  %SYSMAN-I-QUOTA, disk quota statistics on device DUA0: -- Node BARNUM
       UIC                    Usage        Permanent Quota   Overdraft Limit
   [GROUP11,PRIDE]           0            10000             500
```

Notes on Example 6–5:

❶ OPER privilege is required in order to run SYSMAN.

❷ Set the environment to the remote node BARNUM.

❸ SYSPRV privilege is required in order to manipulate disk quotas.

❹ Can't SHOW. Disk quotas are not enabled in the target environment.

❺ Can't ENABLE. QUOTA.SYS file doesn't exist in the target environment.

❻ Create QUOTA.SYS file and automatically enable disk quotas.

❼ Add a disk quota record.

❽ Display a disk quota record.

The DO Command

The DO command executes the specified DCL command or command procedure on all nodes in the current environment. The format for the SYSMAN DO command is:

```
SYSMAN> DO  [/OUTPUT=filespec]  dcl-command-line
```

- Profile must include the privilege(s) required by the DCL command being executed.
- Each DO command creates a new subprocess.
  - No process context retained between DO commands
  - Cannot run a program or command procedure on a remote node if it expects input
  - Must have all parameters on the same command line:

    ```
    SYSMAN> DO @SYS$SYSTEM:SHUTDOWN 10 "Doing backups" NO -
    _SYSMAN> YES "Later" YES CLUSTER_SHUTDOWN
    ```

Example 6–6 SYSMAN DO Command (Single Node)

```
❶ SYSMAN> SET ENVIRONMENT/NODE=BARNUM
%SYSMAN-I-ENV, current command environment:
 Individual nodes: BARNUM
 Username PRIDE       will be used on nonlocal nodes
%SYSMAN-I-PROFRESET, profile on remote nodes has been reset to UAF defaults

❷ SYSMAN> DO SHOW SYSTEM
%SYSMAN-I-OUTPUT, command execution on node BARNUM
❸ VAX/VMS V5.4  on node BARNUM   5-APR-1990 15:05:23.88   Uptime  26 00:26:26
  Pid    Process Name    State  Pri      I/O       CPU         Page flts  Ph.Mem
23200041 SWAPPER         HIB    16         0   0 00:00:35.64          0       0
23200046 CONFIGURE       HIB    10       133   0 00:00:01.90        106     161
23200048 ERRFMT          HIB     7     19908   0 00:07:49.88         82     118
23200049 CACHE_SERVER    HIB    16        60   0 00:00:00.34         62      93
2320004A CLUSTER_SERVER  HIB    10        10   0 00:00:04.81        138     256
2320004B OPCOM           HIB     7     22190   0 00:08:48.10       1220     177
2320004C AUDIT_SERVER    HIB    10      1977   0 00:01:35.23       1776     827
2320004D JOB_CONTROL     HIB    10       757   0 00:00:07.96        159     331
2320004E EVENT_SERVER    HIB     6         8   0 00:00:00.18         65      80
2320004F SMISERVER       HIB     9        59   0 00:00:02.43        331     547
23200053 NETACP          HIB    10     24080   1 06:39:51.86   62856382    1500
23200054 EVL             HIB     6     28012   0 00:13:08.23    1010606      38   N
23200055 REMACP          HIB     8       174   0 00:00:01.40         77      47
23200056 DNS$ADVER       HIB     5     27293   0 00:07:24.79        198     296
23200057 DFS$COM_ACP     HIB    10        12   0 00:00:00.55         97     165
2320049A INSPECT$Exec    HIB     8       573   0 00:00:19.29       2065      55
❹ 2320059D PRIDE_1       CUR     4       225   0 00:00:05.68        658     389
```

Notes on Example 6–6:

❶ Set the environment to a single remote node.

❷ Execute the SHOW SYSTEM command on that node.

❸ Output comes from the remote node.

❹ Current process on the remote node is this user's subprocess.

The DO Command in a VAXcluster Environment

- In a cluster environment, SYSMAN executes the commands sequentially on all nodes in the cluster.
  - Each command executes completely before SYSMAN sends it to the next node in the environment.
  - Any node that is unable to execute the command returns an error message.
  - An error message is displayed if the timeout period expires before the node responds.
- Some DCL commands, such as MOUNT/CLUSTER and SET CLUSTER/EXPECTED_VOTES, operate cluster-wide by design.
  - For these commands to execute successfully in SYSMAN, define the environment to be a single node within the cluster.
- Example 6–7 illustrates the use of the DO command in a VAXcluster environment.

Example 6–7 SYSMAN DO Command (VAXcluster Environment)

```
❶ SYSMAN> SET ENVIRONMENT/CLUSTER

❷ SYSMAN> DO SHOW SYSTEM
  %SYSMAN-I-OUTPUT, command execution on node BAILEY
❸ VAX/VMS V5.4   on node BAILEY  16-APR-1990 14:03:04.80   Uptime  0 01:49:43
    Pid     Process Name     State  Pri      I/O        CPU          Page flts Ph.Mem
  20800021 SWAPPER          HIB     16        0    0 00:00:03.89          0      0
❹ 20800063 DUFFY_1          CUR      4       22    0 00:00:00.42        283    287  S
  20800026 CONFIGURE        HIB     10       33    0 00:00:00.23        109    161
  20800028 ERRFMT           HIB      8      104    0 00:00:00.72         82    118
  20800029 CACHE_SERVER     HIB     16        8    0 00:00:00.08         62     93
  2080002A CLUSTER_SERVER   HIB     10       10    0 00:00:00.32        138    178
  2080002B OPCOM            HIB      7       98    0 00:00:01.34        256    127
  2080002C AUDIT_SERVER     HIB      9       83    0 00:00:01.48       1362    365
  2080002D JOB_CONTROL      HIB      9      171    0 00:00:01.03        165    307
  2080002E EVENT_SERVER     HIB      6        9    0 00:00:00.08         65     80
  2080002F SMISERVER        HIB      9       41    0 00:00:00.86        274    400
  20800031 NETACP           HIB     10       71    0 00:03:16.52      73496   1500
  20800032 REMACP           HIB      8       12    0 00:00:00.13         64     34
  20800033 DNS$ADVER        HIB      4      201    0 00:00:01.57        220    256
  20800034 DFS$COM_ACP      HIB     10       12    0 00:00:00.23         97    164
  20800035 INSPECT$Exec     HIB      8      104    0 00:00:01.54        473     61
  20800038 DUFFY            LEF      5      115    0 00:00:03.95       5326   1722

  %SYSMAN-I-OUTPUT, command execution on node BARNUM

❺ VAX/VMS V5.4   on node BARNUM  16-APR-1990 14:02:57.13   Uptime  0 04:00:37
    Pid     Process Name     State  Pri      I/O        CPU          Page flts Ph.Mem
  20400041 SWAPPER          HIB     16        0    0 00:00:18.64          0      0
  20400046 CONFIGURE        HIB     10       32    0 00:00:00.78        110    158
❻ 20400087 DUFFY_1          CUR      7       29    0 00:00:01.07        279    280  S
  20400048 ERRFMT           HIB      8      199    0 00:00:03.61         82    115
  20400049 CACHE_SERVER     HIB     16        6    0 00:00:00.25         62     90
  2040004A CLUSTER_SERVER   HIB     10       21    0 00:00:01.08        126    214
  2040004B OPCOM            HIB      6      114    0 00:00:04.24        256    133
  2040004C AUDIT_SERVER     HIB     10      112    0 00:00:05.53       1455    278
  2040004D JOB_CONTROL      HIB     10      182    0 00:00:02.16        159    311
  2040004E EVENT_SERVER     HIB      6        6    0 00:00:00.19         62     88
  2040004F SMISERVER        LEF      8       72    0 00:00:02.57        337    478
  20400053 NETACP           HIB     10      134    0 00:11:43.93      77627   1500
  20400054 REMACP           HIB      8       17    0 00:00:00.33         72     45
  20400055 DNS$ADVER        HIB      4      500    0 00:00:10.66        195    297
  20400056 DFS$COM_ACP      HIB     10       11    0 00:00:00.56        112    139
  20400057 INSPECT$Exec     HIB      8       96    0 00:00:04.25        568     59
```

Notes on Example 6–7:

❶ The environment is set to the cluster containing the node BARNUM.

❷ The SHOW SYSTEM command is executed in that environment.

❸ Note that node BAILEY is part of the cluster containing BARNUM.

❹ Current process on the node BAILEY is this user's subprocess.

❺ The next part of the output comes from node BARNUM.

❻ Current process on the node BARNUM is this user's subprocess.

SYSMAN Restrictions

- SYSMAN will execute commands procedures directly.
  - `SYSMAN> @filename`
  - `SYSMAN> DO @filename`

  These will work as long as filename.com does **not** prompt for input.
- SYSMAN does not accept input from a command procedure on a remote system.
- Commands supported cluster-wide should not be executed from SYSMAN cluster-wide.
  - No need to do MOUNT/CLUSTER on each node
- SYSMAN can be run in batch or from a command procedure.
  - It does not read remote passwords if either SYS$COMMAND or SYS$INPUT do not point to a terminal.
  - Remote environments cannot be processed in this way.
- SYSMAN supports key definitions (commands) and initialization files.
- The system and user's login command procedures are **not** executed.
  - The system manager is not logged in individually.
  - Each SYSMAN action is done in the context of the SMISERVER process or, for a DO command, in a subprocess of SMISERVER.

THE LICENSE MANAGEMENT FACILITY (LMF)

LMF allows you to:

- Manage license keys
- INCLUDE or EXCLUDE nodes for subcluster licensing
- Perform administration and tracking
- Provide concurrent user licensing distributed over the cluster
- Centralize license management
- Combine license keys
- Copy freely - it focuses on use
- Manage licenses for the entire VAX family of systems running the VMS operating system

The LICENSE Database

- The disk database is manipulated with the LICENSE command.
- The in-memory database is used when the product (layered or otherwise) makes an inquiry
- Cannot automatically stop software from running
  - — The software queries LMF for information.
  - — If the information sought is not present the software makes the decision to run or not.
- Will not prevent older, existing software, which does not query the database, from running
  - — Over time, Digital software products will use license keys.

INSTALLING LAYERED PRODUCTS ON A COMMON SYSTEM DISK

The standard precautions for product installation hold also for VAXcluster systems. For most products, the only concerns that require more careful attention are SYSGEN parameters and licensing.

- Install layered products as in a non-clustered environment.
- Perform the actual installation (the documented procedure) once for each system disk.

After installation:

- Create product-specific files in the SYS$SPECIFIC directory on each node, if necessary.
  - VMSINSTAL will tell you if you have to create a directory in SYS$SPECIFIC.
- Modify any files in SYS$SPECIFIC on each node that the procedure told you to modify.
- Reboot each node to ensure that:
  - The node is set up to run the product correctly.
  - The node is running the latest version of the product.
- Manually run the Installation Verification Procedure (IVP), if you did not run it during the installation procedure.
  - Run it from at least one other node in the cluster, preferably from all nodes.
  - If VMSINSTAL deletes the IVP, you may be able to restore it with the following command procedure:

```
$ @VMSINSTAL  product  source_device  OPTIONS G  device:[directory]
$ BACKUP  device:[directory]product.A/SELECT=KITINSTAL.COM  device:[directory]
$ @device:[directory]:KITINSTAL  VMI_IVP
```

Do not use search lists when creating files:

- Use SYS$SPECIFIC or SYS$COMMON instead of SYS$SYSROOT
- Use SYS$SPECIFIC:[SYSEXE] or SYS$COMMON:[SYSEXE]) instead of SYS$SYSTEM

VAXcluster OPERATOR COMMUNICATIONS

Enabling and Disabling Operator Terminals

State transition messages:

- In large clusters, state transitions will generate a large number of multi-line OPCOM messages as nodes join and leave the cluster.
- To eliminate such messages, you can do one of the following:
  - — Include the following DCL commands in the site-specific startup command file, SYSTARTUP_V5.COM:

    ```
    $ DEFINE/USER SYS$COMMAND OPA0:
    $ REPLY/DISABLE=CLUSTER
    ```

  - — Enter the above command interactively from the system manager's account.

To disable the console terminal at startup time:

- Place these commands in SYSTARTUP_V5.COM:

```
$ DEFINE/USER SYS$COMMAND OPA0:
$ REPLY/DISABLE
```

To enable another terminal (for example, BAILEY$TTA0) at startup time:

- Place these commands in SYSTARTUP_V5.COM:

```
$ DEFINE/USER SYS$COMMAND BAILEY$TTA0:
$ REPLY/ENABLE
```

Using the REPLY Command to Communicate with Users

- **To reply to all users on all nodes:**

```
$ REPLY /BELL /USER -
"Warning: BAILEY may crash at any time"
```

- **To reply to all users on your own node:**

```
$ REPLY /BELL /USER /NODE -
"Who left the tape on MFA0:?"
```

- **To reply to all users on certain nodes:**

```
$ REPLY /BELL /USER /NODE=(BARNUM,BAILEY) -
"Who left the tape on MFA0:?"
```

- **To reply to specific users on all nodes:**

```
$ REPLY /BELL /USER=(PAVAROTTI, DOMINGO) -
"Please pick up your listings."
```

- **To reply to specific users on specific nodes:**

```
$ REPLY /BELL /USER=(PAVAROTTI, DOMINGO) /NODE=BARNUM -
"Please pick up your listings."
```

- **To reply to specific terminals on specific nodes:**

```
$ REPLY /TERMINAL=(BARNUM$TTA0,BAILEY$TTA0) -
"Please log off."
```

RECONFIGURING THE CLUSTER

The system manager can, if need arises, change the hardware configuration of the cluster. Some of the changes you can make in a running cluster include the following:

- You can place a disk off-line, and later place it back on-line. You can even unplug a DSA disk from its controller.
- You can place an HSC subsystem off-line (and even remove it from the cluster hardware configuration) and later return it to the cluster. If HSC disks are dual-ported, you can remove the HSC without affecting users.
- With proper control of the expected-votes and cluster quorum, you can remove a VAX node from, or add a VAX node to the VAXcluster system.
- You can shut down the entire cluster at once.

Deconfiguring a Local Unserved Disk

- Ensure that all files are closed
- Dismount the disk
  - Prevents hung processes
  - Prevents the need to rebuild volumes
- There are special considerations for deconfiguring the following:
  - MSCP served disks
  - HSC disks

disk weghalen voor het hele cluster:
dismount /cluster/abort

Deconfiguring an MSCP Served Disk

- Check for open files on all members with access to the disk.
- Dismount the disk from every system that has it mounted.
  - — Use the DISMOUNT/CLUSTER command. /ABort
  - — Use SHOW DEVICE on all nodes (from SYSMAN) to make sure the disk is no longer mounted.
- If any system boots, pages, or swaps using this disk, shut down the system. If possible, the page and swap files may be DEINSTALLED in SYSGEN and migrated to another page or swap file location, unless the file is the primary paging file or the only paging file.
- Take the disk drive off-line by setting its port select switches to the out (off-line) position.

Deconfiguring an HSC Disk

- Check for open files on all members with access to the disk.
- Dismount the disk from every system that has it mounted.
  - — Use the DISMOUNT/CLUSTER command.
- If any system boots, pages, or swaps using this disk, shut down the system. If possible, the page and swap files may be DEINSTALLED in SYSGEN and migrated to another page or swap file location.
- Issue a SHOW DISKS command at the HSC console.
  - — If the disk is dual-ported, check both HSC units.
  - — The disk should not appear on-line.
  - — The port select lights on the disk drive should not be lit.
- Take the disk drive off-line by setting its port select switches to the out (off-line) position.

Reconfiguring HSC units

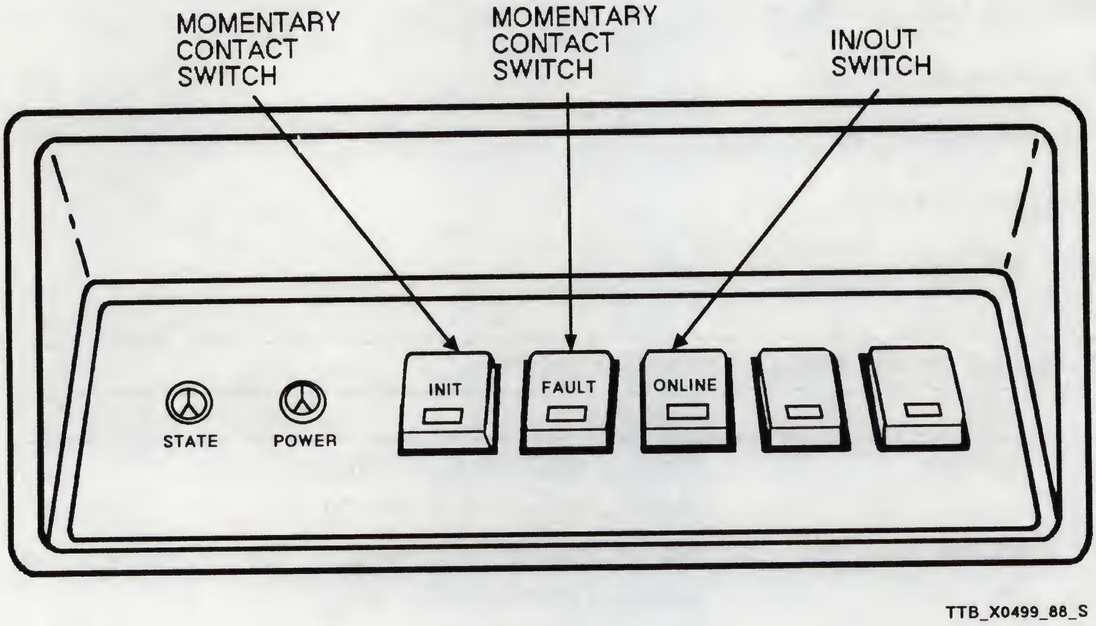
You can control the HSC unit from:

- The HSC console terminal
- A remote terminal ($SET HOST /HSC)
- The operator control panel
- HSC interior switches

Table 6–4 Indicators on the HSC Operator Control Panel

Indicator	State	Meaning
STATE	Blinking	Slowly: HSC unit is operating normally Rapidly: HSC unit is booting.
	Steady on or off	HSC is unit halted or hung.
POWER	On	Power is at the correct level.
INIT	On	HSC unit is booting.
FAULT	On	HSC unit has detected a fatal error and has halted.
ONLINE	Off	No member has a virtual circuit open to the HSC unit.
	On	A member has established a virtual circuit to the HSC unit.
Unlabeled		(Used in fault code display.)

Figure 6–1 HSC Operator Control Panel



Removing an HSC unit

To place an HSC unit off-line:

- Run SHOW ALL on the HSC to determine the current parameter settings.
- Set the ONLINE switch on the HSC front panel to the out position.
  - This action prevents other systems from connecting to it, but does not close existing connections.
- Dismount and place off-line each tape drive on this HSC unit.
- Dismount and place off-line each single-ported disk on this HSC unit
  - Dismount it from every system that has it mounted.
  - Shut down every system that boots from or pages/swaps to it.
  - Set its port select switches to the out position.
- Cause failover of each dual-ported disk on this HSC unit
  - The port select switch for the other HSC unit should be in the in position.
  - Set the port select switch for this HSC unit to the out position.
  - The port select light for this HSC unit will go off.
  - Disk I/O will continue, using the other HSC unit
  - Failover does not occur for disks mounted /FOREIGN (they go off-line instead).
- Issue the SHOW DISKS command at the HSC console terminal.
  - No disks should appear as on-line.

Returning an HSC to the Cluster

When returning an HSC to the cluster:

- Make sure its parameters are set the same as they were before.
- If you bring it on-line with a different NAME, ID, or allocation class, you will have to reboot all of the VAX systems.

To return the HSC to the cluster:

- Make sure the ONLINE switch is in the OFFLINE (out) position.
- Set the SECURE/ENABLE switch to the ENABLE position.
- Set the disks' PORT SELECT switches for this HSC unit in the OFFLINE position.
- Type CTRL/Y, then RUN SETSHO.
- Use the SHOW SYS command to check parameter settings.
- Use the SET command to change parameters if necessary.
- When you EXIT from SETSHO, the HSC unit reboots automatically, if you have changed any parameters.
- After the HSC reboots, press the ONLINE switch and make disks available by pushing their port select switches.
- Set the SECURE/ENABLE switch to the SECURE position.

Reconfiguring VMS Members

Adding or removing VMS members requires attention to three issues:

- The number of votes (and the value of EXPECTED_VOTES) in the cluster. This determines whether the cluster will be able to maintain or achieve a quorum.
- The presence of a mutually agreed upon QUORUM DISK whose votes can be counted toward achieving quorum.
- The value of LOCKDIRWT.

Assuming that quorum can be achieved, a non-zero value for LOCKDIRWT will require a partial rebuild of the lock manager database, specifically to add or remove the node from the lock directory vector and adjust the database accordingly. If a quorum disk is specified, there will be a delay, during the cluster state transition, while all nodes confirm that they agree on the quorum disk and can all read and write to the QUORUM.DAT file, before counting the quorum disk's votes. This delay, at maximum, is four times the QDSKINTERVAL (which should be set the same on all nodes that are quorum disk watchers).

Shutting Down in the VAXcluster Environment

There are four options for shutting down cluster nodes:

- **REMOVE_NODE**
  - For nodes that will not rejoin for an extended time
  - Active quorum is adjusted downwards, if possible
- **CLUSTER_SHUTDOWN**
  - Coordinated shutdown of all nodes
  - Nodes suspend activity just short of complete shutdown
  - When all nodes reach this level, they all shut down together
- **REBOOT_CHECK**
  - VMS system files necessary for reboot are checked

    Notification of any missing files occurs

    Files must be replaced before proceeding

- **SAVE_FEEDBACK**
  - — Saves the AUTOGEN FEEDBACK information
  - — Makes a copy of the workload data from memory and places it in a file
  - — After reboot, this file can be used to reset system parameters.
  - — Do not use this option when the system workload prior to the last reboot has been atypical.

Removing a VAX Node Temporarily

- Check to see which nodes are up, and count how many votes they have.

```
$ SHOW CLUSTER/CONTINUOUS
COMMAND> ADD VOTES,CL_QUORUM,CL_VOTES
```

- Dismount, cluster-wide, all MSCP served disks local to the node.
- Shut down the system, requesting the REMOVE_NODE option.

```
$ @SYS$SYSTEM:SHUTDOWN.COM
```

- When you reboot the system, the cluster quorum will return to its original value.

While the system is down:

- The cluster may be vulnerable.
  - — If another node fails, the remaining nodes may lose quorum.
  - — If another node reboots, the cluster quorum value may increase to its normal value.

Removing a VAX Node Permanently

- Check to see which nodes are up, and count how many votes they have.

```
$ SHOW CLUSTER/CONTINUOUS
COMMAND> ADD VOTES, CL_QUORUM, CL_VOTES
```

- Dismount, cluster-wide, all MSCP served disks local to the node.
- Shut down the system, requesting the REMOVE_NODE option.

```
$ @SYSSYSTEM:SHUTDOWN.COM
```

- After the system is down:
  - — Enter the command SET CLUSTER/EXPECTED_VOTES to set the current cluster_ quorum value equal to (the votes now in the cluster + 2)/2.
  - — If you lose quorum and the VAXcluster system hangs, invoke IPC and enter the following commands at one of the nodes:

```
CTRL/P
>>>HALT
>>>D/I 14 C
>>>CONT
IPC> Q
```

Store in reg. 20 12 (forceer state transitie)

Daarna gaat het cluster verder

The Q command recalculates the quorum on the cluster

- On the remaining systems:
  - — Set the SYSGEN parameter EXPECTED_VOTES to the total number of votes now in the cluster.
  - — This procedure prevents the cluster expected-votes value from going back to its original value if a system reboots.
- If you return the node to the cluster, remember to return the SYSGEN parameter EXPECTED_VOTES to its normal value on every node.
- While the node is down, the cluster can survive the failure of other nodes.

For satellites, and other nodes with zero (0) votes, these concerns do not apply.

RESTORING CLUSTER QUORUM

To display the current cluster votes information, enter the following commands:

```
$ SET CLUSTER/EXPECTED_VOTES
```

This shows the current cluster_quorum and cluster_votes.

To look at individual node SYSGEN parameter settings:

```
$ MCR SYSMAN SET ENVIRONMENT/CLUSTER
SYSMAN> PARAMETER SHOW EXPECTED_VOTES
SYSMAN> PARAMETER SHOW VOTES
```

Reducing Cluster Quorum

To reduce the cluster quorum value:

- Enter the following command on any one node in the cluster:

  ```
  $ SET CLUSTER/EXPECTED_VOTES=n
  ```

- The command to change cluster quorum to 2 is:

  ```
  $ SET CLUSTER/EXPECTED_VOTES=3
  ```

- Adjust EXPECTED_VOTES parameter in MODPARAMS.DAT

Normal Shutdown of a VAXcluster System

To perform a normal shutdown of the entire cluster:

- On each node, enter @SYS$SYSTEM:SHUTDOWN.
- Request the CLUSTER_SHUTDOWN option.
  - If you do not request CLUSTER_SHUTDOWN, the cluster will lose quorum and hang after a certain number of nodes have shut down.
  - Use CLUSTER_SHUTDOWN only when you are shutting down all nodes.
- When all of the systems reach a certain point in the procedure, they shut down together.

Using MODPARAMS.DAT to Modify SYSGEN Parameters

To display VAXcluster related SYSGEN parameters:

```
$ RUN SYS$SYSTEM:SYSGEN
SYSGEN> SHOW /CLUSTER
SYSGEN> SHOW /SCS
```

To modify SYSGEN parameters:

- Edit SYS$SPECIFIC:[SYSEXE]MODPARAMS.DAT
- Execute SYS$UPDATE:AUTOGEN.COM

Advantages of using AUTOGEN rather than SYSGEN:

- AUTOGEN uses a feedback mechanism to set parameter values based on your system's workload.
- AUTOGEN reconfigures other parameters to reflect your changes.
- You have a record of changes in MODPARAMS.DAT.
- Changes recorded in MODPARAMS.DAT are not lost during VMS updates.

VAXcluster BACKUP OPERATIONS

Backup operations that are the **same** in clusters and single VAX systems:

- Backup of a disk while it is in use:
  - For example, incremental backup
- Backup of a nonshared disk

Shadow backup
HSC backup

Backup operations that are **different** in clusters and single VAX systems:

- Protecting a shared disk from access during backup:
  - For example, during image backup
- Image backup of a system disk
  - On a running system
  - From another system
  - Using standalone BACKUP

Backup Tools in a Cluster

Backup tools available for use in a cluster:

- On-line BACKUP
- Standalone BACKUP
- The HSC BACKUP utility

On-Line BACKUP

- Is used from a running system
- Backs up:
  - — Local disks
  - — Cluster-available nonsystem disks
  - — The system disk (files open for writing may not be backed up correctly)

Standalone BACKUP

- Is used to back up a system disk
- Should be used with caution since it does not:
  - — Participate in the cluster
  - — Synchronize volume ownership or file I/O with other cluster systems
- Boots with SCSNODE = "SABKUP", SCSSYSTEMID = 0
  - — Remains as a BRK_NON system in SHOW CLUSTER display
- Can boot from the system disk instead of from console media
  - — Standalone BACKUP is built in the reserved root on any system disk, [SYSE].
  - — @SYS$UPDATE:STABACKIT.COM builds standalone BACKUP in [SYSE] (or in [SYS0] if the disk is not a system disk)

HSC BACKUP and RESTORE

The HSC BACKUP and RESTORE utilities can be used to:

- Back up an HSC disk to HSC tapes
- Restore an HSC disk from an HSC tape backup save set

The HSC BACKUP and RESTORE operations:

- Perform a fast physical backup
- Require that the disk be dismounted
- Cannot be used for incremental backups

- Shadowing Backup (= fysical backup)

schrijven naar een shadow-set alleen naar de master
lezen van " " " van alle schijven

schijven moeten identiek zijn
quorum disk kan niet in een shadow_set.

sysgen par. shadowing 1 Controller Based (alleen met HSC)
2 Host Based (zonder HSC)
3 Beide

Sample Backup Operations

Image Backup of a Shared Nonsystem Disk

The following steps apply to either an HSC disk or an MSCP served disk:

- Dismount the disk from every system that has it mounted.

```
$ DISMOUNT/CLUSTER device
$ SHOW DEVICE device
```

- Mount the disk privately on the system that is to perform the backup.
- Back up the disk.
- Dismount the disk.
- Mount the disk on every system that normally has it mounted.

Example 6–8 Backup of a Shared Nonsystem Disk

```
! Dismount the disk. Make sure it is dismounted on all systems.
!
$ DISMOUNT/NOUNLOAD/CLUSTER $1$DUA1:
$ SHOW DEVICE $1$DUA1:

Device                  Device         Error    Volume
Name                    Status         Count     Label
$1$DUA1:              Online            0

$!
$! Mount it privately
$!
$ MOUNT $1$DUA1: VOLUME1
$!
$! Back it up
$!
$ MOUNT/FOREIGN MFA0:
$ BACKUP/IMAGE/RECORD/IGNORE=LABEL_PROCESSING $1$DUA1: MFA0:USERDISK.SAV
$ DISMOUNT MFA0:
$!
$! Dismount the disk again
$!
$ DISMOUNT/NOUNLOAD $1$DUA1:
$!
$! Mount it cluster-wide
$!
$ MOUNT/CLUSTER $1$DUA1: VOLUME1
%MOUNT-I-MOUNTED, VOLUME1    mounted on _$1$DUA1:
```

Image Backup of a System Disk

There are three ways to do an image backup of a system disk:

- Back up the system disk from the running system
- Boot the system from an alternate system disk and perform a backup of the original system disk.
- Boot standalone backup and perform the backup operation.
- Table 6–5 compares these methods.

Table 6–5 Image Backup of a System Disk

Method	Advantages	Disadvantages
From a running system (HSC or non HSC disk)	Systems can remain up	Backup may be incomplete or inconsistent
From another system	Backup is complete	There must be more than one system disk in the cluster
Standalone backup	Backup is complete	Requires caution

Image Backup of the System Disk on a Running System

- The system disk cannot be write-locked.
- By default, files open for writing are not backed up.
  - BACKUP/IGNORE=INTERLOCK allows backup of files open for writing.
  - Some files within the save set may not be usable.
- Keep system disk activity to a minimum (if possible).
  - Disable logins and stop users.
  - Stop queues.
- Files that are not correctly backed up may include:
  - The current operator log, OPERATOR.LOG
  - The queue file, JBCSYSQUE.DAT
  - The current error log, ERRLOG.SYS
  - The current accounting file, ACCOUNTNG.DAT
  - The Mail database, VMSMAIL_PROFILE.DATA
  - Any other file that is open for writing
- Warnings appear for files open for write access (from any node)
- Example 6–9 shows three files that are commonly not correctly backed up in an image backup of a running system.

Example 6–9 Backup of a System Disk on a Running System

```
$ MOUNT/FOREIGN MFA0:
$ BACKUP/IMAGE/IGNORE=(INTERLOCK, LABEL) SYS$SYSDEVICE: -
_$ MFA0:SYSDISK.SAV
%BACKUP-W-ACCONFLICT, SYS$SYSDEVICE:[SYS0.SYSMGR]ACCOUNTNG.DAT;430
is open for write by another user
%BACKUP-W-ACCONFLICT, SYS$SYSDEVICE:[SYS0.SYSMGR]OPERATOR.LOG;21
is open for write by another user
%BACKUP-W-ACCONFLICT, SYS$SYSDEVICE:[SYS0.SYSEXE]VMSMAIL_PROFILE.DATA;3
is open for write by another user
$ DISMOUNT MFA0:
```

Image Backup of an HSC Connected System Disk from Another System

If there is more than one system disk in the cluster, you may treat the system disk as a nonsystem disk on another system.

The following steps apply to both private and common system disks:

- Shut down every system that boots from the disk.
- On one of the other systems, mount the disk privately, and back it up.
- Reboot the systems that were shut down.

Using Standalone BACKUP (HSC or Non HSC Disk)

To guarantee a consistent backup of an HSC disk:

- Shut down every system that boots from the disk.

  (or)

- Dismount the disk from every system that has it mounted.

Procedure for standalone BACKUP:

- Shut down the system you will be using for standalone BACKUP.
- Boot standalone BACKUP on that system.
- If the disk you are backing up is on an HSC unit, shut down every system that boots from the disk and dismount it from every system that has it mounted.
  - Note that if the disk is local, it is protected from use by the rest of the cluster because standalone BACKUP has no MSCP server.
- If the target device is a cluster-available disk, use SHOW DEVICE to verify that it is not mounted on any system.
  - Make sure no other system mounts the target device during the backup.
- Back up the disk using standalone BACKUP.
- Reboot the VMS operating system on the standalone machine.
- Reboot any other systems you have shut down.
- Mount the disk on every system that normally has it mounted.
  - This step is unnecessary if the startup procedure contains MOUNT/CLUSTER.

Using VAX Volume Shadowing Software as a Supplement to BACKUP

VAX Volume Shadowing software can be used to create a physical backup of an HSC volume.

This example assumes Phase I (controller-based) volume shadowing involving two disks on-line to the same HSC unit.

To use volume shadowing to back up an HSC volume:

- Create the shadow set.

label van DUA3

```
$ MOUNT/SYSTEM $1$DUS0:/SHADOW=$1$DUA3: MEDICVOL
```

This command creates a shadow set with one member volume.

- Add the disk to be used as the backup.

```
$ MOUNT/SYSTEM $1$DUS0:/SHADOW=$1$DUA4: MEDICVOL
```

This command adds the second member to the shadow set with a copy operation.

- Wait for the copy operation to complete. Messages appear at any operator console.
- Dismount the shadow set.

```
$ DISMOUNT $1$DUS0:
```

This command dismounts the shadow set.

- 1DUA4 is now a physical backup of 1DUA3.
- Remount the shadow set without 1DUA4 or 1DUA3.
- Note that the shadow disk was unavailable for a very short period of time.

Phase II (host-based) volume shadowing permits the same operations through VMS hosts. The only differences are in the potential shadow names (DSAnnnn:) and physical device names (which need not be in the same allocation class).

UPDATING VMS IN A VAXcluster SYSTEM

When updating to a new version of the VMS operating system:

- Perform the update procedure once for each system disk.
  - The procedure for a private system disk is the same as for a single system.
  - The procedure will be in that version's release notes.
- After updating a common system disk, run AUTOGEN on each node that boots from that disk.

Rolling Upgrade

For details of the rolling upgrade and specific version of the VMS operating system, see the appropriate Release Notes.

- Allows a cluster to have members that are one release of VMS software apart

  (For example, VMS V5.3 and VMS 5.4 can both be running on different members in the same cluster.)
- All systems booting from the same system disk **must** be running the same version.
- System files that change formats and enhance performance must stay in the old format until all cluster members have been upgraded.
- Mixed-version clusters are intended to be a temporary state, to allow a cluster to continue running throughout even a complex upgrade.

BASIC SECURITY CONSIDERATIONS

Since many local area and mixed-interconnect VAXcluster systems can exist on the same Ethernet, security measures have been developed to protect clusters from accidental or deliberate breach.

Security features available include:

- Cluster authorization file, SYS$COMMON:[SYSEXE]CLUSTER_AUTHORIZE.DAT
  - Updated by SYSMAN
  - Includes cluster group number and cluster password
  - Group or password change, reboot cluster
- Control of conversational bootstrap operations on satellites
  - Default for NISCS_CONV_BOOT is 0 (disabled)
  - Use default value
  - Enable for particular nodes if you wish later

    Parameters exist in device:[SYSx.SYSEXE]VAXVMSSYS.PAR

    Set parameter to 1

    For example, to enable conversational boot from root 10 on device 1DJA11:

```
$ RUN SYS$SYSTEM:SYSGEN
SYSGEN> USE $1$DJA11:[SYS10.SYSEXE]VAXVMSSYS.PAR
SYSGEN> SET NISCS_CONV_BOOT 1
SYSGEN> WRITE $1$DJA11:[SYS10.SYSEXE]VAXVMSSYS.PAR
SYSGEN> EXIT
$
```

Digital strongly recommends that all site security administrators take the following steps:

- Disable (or remove) all Digital default accounts except SYSTEM in any active SYSUAF.DAT unless you have a specific need for these accounts.
- Consider using the password generator for all privileged accounts.
- Review all network proxies and eliminate, if possible, any proxies into privileged accounts.
- Remove the default DECnet account and substitute separate accounts for each network object required for a particular node. NETCONFIG.COM will set this up by default, since VMS version 5.2.

SUMMARY

Typical cluster management tasks include:

- Using the SYSMAN utility
- Managing licenses and license keys in the cluster
- Installing layered products
- Performing operator communications
- Removing cluster-wide disks without disrupting the cluster
- Removing and returning HSC units
- Removing and returning active nodes
- Performing backup operations
- Updating the operating system
- Setting system time
- Restoring cluster quorum after node failure
- Executing conditional shutdown operations
- Performing security functions specific to a local area or mixed-interconnect VAXcluster system

Locating VAXcluster Problems

INTRODUCTION

To locate and solve problems without disrupting operations, a system manager must be familiar with the VAXcluster configuration. This module discusses a number of VMS tools that can provide the system manager with information about the configuration of a VAXcluster system. It will also show you how to interpret certain diagnostic information and how to deal with some problems when they occur.

OBJECTIVES

After completing this module, students should know how to:

- Deal with a node failure
- Deal with the failure of a node to join a VAXcluster system
- Interpret VAXcluster system-specific error messages
- Use a variety of utilities to gather diagnostic information relating to VAXcluster system problems

RESOURCES

- *Guide to Maintaining a VMS System*
- *VMS VAXcluster Manual*
- *VMS Show Cluster Utility Manual*
- *VMS Monitor Utility Manual*
- *VMS Networking Manual*
- *VMS Network Control Program Manual*
- *VMS System Generation Utility Manual*
- *VMS System Dump Analyzer Utility Manual*
- *HSC User Guide*
- *VAXsimPLUS User Guide*
- *Getting Started with VAXsimPLUS*

FAILURE OF NODES TO BOOT OR JOIN THE CLUSTER

Normal Cluster Node Boot

To effectively diagnose a problem, system managers should know the events that occur during a normal boot procedure of a cluster node.

1. The node boots.
2. If the node finds a cluster, the node attempts to join and the CONNECTION MANAGER displays messages.
3. If no cluster is found, the CONNECTION MANAGER forms a cluster when enough nodes have booted to establish quorum.
4. If quorum is lost while the node is booting, or if the node is unable to join within two minutes, %CNXMAN displays messages showing connections made.
5. Once the node has joined, normal startup procedures execute.

There are many diagnostic messages that are useful to the system manager if this procedure fails.

- Connection manager (%CNXMAN) messages
- VAXport error messages (Ethernet of CI of DSSI)
- HSC error messages and fault conditions
- Error log entries
- DECnet network event messages

Failure of CI Node to Boot

If a CI node fails to boot:

- Check that SCSNODE and S~~Y~~CSSYSTEMID parameters are unique
- Check bootstrap command file
  - — HSC node number
  - — HSC disk unit number
  - — Root number

  DEFBOO.CMD
  SHOW BOOT FLAG
  #de pollen poorten

- Check that SYSGEN parameter PAMAXPORT is equal to or greater than the highest CI port number.
- Is the HSC unit on-line?
- Check that disk is available show disk of VTDPY
- Check that node has access to HSC unit
  - — SHOW HOSTS
  - — HSC SETSHO

HSC Diagnostic Messages and Fault Codes

Diagnostic messages appear at the HSC console terminal.

- If FAULT light is on, push and release the FAULT switch; the lights will display a pattern indicating a certain fault.
- If the fault is TU58 drive FAILURE — HSC50 — or MISSING REQUIRED FILES:
  - — Make sure system tape is in TU58 drive 0. on write enabled.
  - — Try the backup system tape.
- For other faults:
  - — Write down the fault code. (Refer to *HSC User Guide*.)
  - — Try rebooting the HSC unit.
  - — Call your Customer Service representative.

If your HSC disks are dual-ported, an HSC unit may fail without affecting the rest of the cluster.

- Disks automatically fail over if they are not mounted /FOREIGN, so you may not notice that the HSC has failed.
- Diagnostics are especially important.
  - — Look at the indicator lights and console printout regularly.

HSC SETSHO Utility

The SETSHO utility can be used as a diagnostic tool.

To invoke the HSC SETSHO utility, enter one of these commands at the HSC prompt:

- SHOW ALL (to display all parameters)
- SHOW parameter (to show the given parameter)
- SET parameter (to set the given parameter)
- RUN SETSHO (to enter interactive SETSHO)
- For a list of available parameters:
  - At the HSC prompt, type RUN SETSHO
  - At the SETSHO> prompt, enter HELP

While your HSC units are running correctly:

- Enter the command SHOW ALL on each HSC unit
- Keep the printouts for later reference
- Important things to look for:
  - Name and system ID
  - Allocation class
  - Status of both disks and tapes
  - Virtual circuits
  - Revision levels

Example 7-1 HSC SHOW Output (Sheet 1 of 3)

```
HSC50> SHOW ALL

 9-Jun-1989 16:39:25.32  Boot: 28-May-1989 13:02:19.32  Up: 291:37
Version: V350            System ID: %X000000000003      Name: CLOWN
Front Panel-secured      Sector size-512
Console Dump Enabled     Load Device Dump Disabled
Restart - Warm
Automatic Diagnostics  Enabled
Periodic Diagnostic Interval- 1  Enabled
DISK allocation class =   1    TAPE allocation class =    5
Start command file  Disabled
- - - - - - - - - - - - - - - -
Current modules that will be loaded:

        Central Error Logging
        Demon
        DUP
- - - - - - - - - - - - - - - -
Error Levels Displayed Next Reinitialization:

        Error
        Fatal
Current Error Levels Displayed:

        Error
        Fatal
- - - - - - - - - - - - - - - -
Outband Levels Displayed Next Reinitialization:

        Error
        Fatal
Current Outband Levels Displayed:

        Error
        Fatal
- - - - - - - - - - - - - - - -
Current ODT setting:

- - - - - - - - - - - - - - - -
```

Example 7–1 HSC SHOW Output (Sheet 2 of 3)

```
Disabled/Enabled Hosts (Node Number):

        Node     Status
        ----     ------
        0        Enabled
        1        Enabled
        2        Enabled
        3        Enabled
        4        Enabled
        5        Enabled
        6        Enabled
        7        Enabled
        8        Enabled
        9        Enabled
        10       Enabled
        11       Enabled
        12       Enabled
        13       Enabled
        14       Enabled
        15       Enabled
- - - - - - - - - - - - - - - - -
R #      Status          Type            Version
---      ------          ----            -------
9        Enabled         Req. empty
8        Enabled         Req. empty
7        Enabled         K.sti           MC- 27 DS- 2
6        Enabled         Req. empty
5        Enabled         Req. empty
4        Enabled         Req. empty
3        Enabled         K.sdi           MC- 40 DS- 3
2        Enabled         K.sdi           MC- 40 DS- 3
1        Enabled         K.ci            MC- 43 DS- 2 Pila-65 K.pli-5
0        Enabled         P.ioc
- - - - - - - - - - - - - - - - -
Unit #    R #   Port    Type    State  /  Version
------    ---   ----    ----    -----------------
   0      2      1      RA82    online - host access / MC-  8 HV-  8
   1      3      2      RA82    online - host access / MC-  7 HV-  6

Drives stored in saved NOHOST table
SETSHO-I The NOHOST_ACCESS table is empty.
- - - - - - - - - - - - - - - - -
Unit #    R #   Port    Type    State  /  Version
------    ---   ----    ----    -----------------
    0      7      0     TA79    unavailable/offline  / DMC-  0 DHV-  0 FMC-  4 F
HV-255

Drives stored in saved NOHOST table
SETSHO-I The NOHOST_ACCESS table is empty.
- - - - - - - - - - - - - - - - -
```

Example 7–1 HSC SHOW Output (Sheet 3 of 3)

```
Maximum Number of Tapes - 24
- - - - - - - - - - - - - - - -
Maximum Number of Formatters - 24
- - - - - - - - - - - - - - - -
Statistics for node number ***all***

  Path A  Path B
ACKs:  3596272  3593004
NAKs:  576490  576611
No Responses 1508383  1508573

Logging now enabled for node number ***all***
- - - - - - - - - - - - - - - -
Virtual circuits are open to:

Node 0  Path A Path B
Node 1  Path A Path B
2 - Available Nodes
- - - - - - - - - - - - - - - -
 Free Lists
1131 - Control Blocks
  32 - Short Lifetime Control Blocks
 212 - Free Buffers
- - - - - - - - - - - - - - - -
Disabled Memory List:

 None

Suspect Memory List:

 None

MAX MEMORY AVAILABLE

Program Memory:  131072 words
Control Memory:  65536 words
Data Memory:  65536 words

ACTUAL MEMORY AVAILABLE

Program Memory:  131072 words
Control Memory:  65536 words
Data Memory:  65536 words

- - - - - - - - - - - - - - -
Line - 0 Terminal  noscope Enabled
Line - 2 Terminal  scope Enabled
Line - 4 Terminal  noscope Enabled
Line - 5 Terminal  noscope Enabled
Line - 0 Reserved
Line - 1 Reserved
Line - 2 Reserved
Line - 3 Reserved

SETSHO Program Exit
```

Node Fails to Join the Cluster

If a node boots, but fails to join the cluster:

- Check that VAXcluster software has been loaded. Vaxcluster op 1 of 2?
- Check for the existence of the cluster security database file, SYS$COMMON:CLUSTER_AUTHORIZE.DAT. groupnumber + password juist
- Check that node booted from correct disk and system root.
- Check that SCSNODE and SCSSYSTEMID are unique.

Satellite Node Fails to Boot

If a satellite node fails to boot:

- Log in as SYSTEM on the boot server. onder privileged account
- If event logging is not enabled, enter the following NCP command:

```
NCP> SET LOGGING MONITOR EVENT 0.*
NCP> SET LOGGING MONITOR STATE ON
```

- Enter the following DCL command:

```
$ REPLY/ENABLE=NETWORK
```

- Boot the satellite
  - — If no messages appear, there may be problems with physical cable connections or adapter service.
  - — The MOP service request message looks like this:

```
DECnet event 0.3, automatic line service
From node 3.1 (WHYNOT), 24-MAY-1989 10:34:24.01
Circuit QNA-0, Load, Requested, Node = 3.12 (BECAUS)
File = SYS$SYSDEVICE:<SYS10.>, Operating system
Ethernet address = 08-00-2B-17-AC-04
```

  - — If information in the DECnet database is incorrect, a message displaying the correct address will appear.

Komt via ethernet met het request mee

```
DECnet event 0.7, aborted service request
From node 3.1 (WHYNOT), 24-MAY-1989 11:07:21.18
Circuit QNA-0, Line open error, Ethernet address = 08-00-2B-17-AC-04
```

echte adres van de Satelliet. In de database eventueel met sys$manager: cluster_config change veranderen

- If the node fails to boot
  - — Check that the Ethernet hardware address is correct.
  - — Check that boot device is available.
  - — Check that DECnet software is up and running.
  - — Check that circuit service is enabled.

```
NCP> SHOW CIRCUIT circuit-id
```

  - — To enable circuit service:

```
NCP> SET CIRCUIT circuit-id STATE OFF
NCP> DEFINE CIRCUIT circuit-id SERVICE ENABLED
NCP> SET CIRCUIT circuit-id SERVICE ENABLED STATE ON
```

  - — Check that the Ethernet hardware address is correct.

Using NCP in a VAXcluster System

The network control program (NCP) displays information about:

- Nodes
- Circuits
- Lines
- The network

Common NCP commands:

- SHOW KNOWN NODE CHARACTERISTICS

  Information about all currently defined DECnet nodes
- SHOW KNOWN CIRCUIT CHARACTERISTICS

  Information about all currently defined DECnet circuits
- SHOW KNOWN LINE CHARACTERISTICS

  Information about all currently defined DECnet lines

SHOW KNOWN NODE CHARACTERISTICS

Example 7–2 Output from SHOW KNOWN NODE CHARACTERISTICS

```
Known Node Volatile Characteristics as of 27-MAY-1990 23:53:51

Executor node =  1.1   (BARNUM)

State                      = on
Identification             = DECnet VAX V5.4,  VMS V5.4
❶
      .
      .
      .
❷ Remote node =    1.99  (LION)

❸ Hardware address         = 08-00-33-41-77-9F
Tertiary loader            = SYS$SYSTEM:TERTIARY_VMB.EXE
Load Assist Agent          = SYS$SHARE:NISCS_LAA.EXE
Load Assist Parameter      = BARNUM$DJA0:<SYS10.>
      .
      .
      .
```

Notes on Example 7–2:

❶ In actual usage, SHOW KNOWN NODES would display additional information about the executor node. In this example the executor node information has been condensed.

❷ The VAX system BARNUM serves as a boot node for a MicroVAX node LION.

❸ The parameters defined in the DECnet database allow BARNUM to initiate the cluster system boot process for the node LION.

SHOW KNOWN CIRCUIT CHARACTERISTICS

Example 7–3 Output from SHOW KNOWN CIRCUIT CHARACTERISTICS

```
Known Circuit Volatile Characteristics as of 27-MAY-1990 22:36:15

❶ Circuit = QNA-0

❷ State                   = on
❸ Service                 = enabled
  Designated router       = 1.1 (BARNUM)
  Cost                    = 4
  Router priority         = 64
  Maximum routers allowed = 33
  Hello timer             = 15
  Type                    = Ethernet
  Adjacent node           = 1.1 (BARNUM)
  Listen timer            = 90
```

Notes on Example 7–3:

❶ The circuit field displays the name of the circuit as defined in the DECnet database.

❷ The state field is on. This circuit is available to process network traffic.

❸ The service field must be set to enabled on the boot node for downline loading to function correctly.

SHOW KNOWN LINE CHARACTERISTICS

Example 7–4 Output from SHOW KNOWN LINE CHARACTERISTICS

```
Known Line Volatile Characteristics as of 28-MAY-1990 00:38:27

Line = QNA-0

Receive buffers          = 6
Controller               = normal
Protocol                 = Ethernet
Service timer            = 4000
Hardware address         = 08-00-33-41-77-9F
Device buffer size       = 1498
```

van eigen node

Connection Manager (%CNXMAN) Messages

- Provide information about VAXcluster system transitions
- Displayed at the console terminal
- Content, not order of messages important
- When OPCOM is running, similar OPCOM messages appear:
  - On operator terminals (if enabled as CLUSTER operator)
  - In the operator log

Know what is normal for your cluster:

- The more familiar you are with the behavior of a healthy cluster, the easier it will be for you to diagnose problems.
- Become familiar with the messages displayed when your cluster is running normally.
- This allows you to recognize abnormal messages when they appear.
- Examine console output regularly, since there may be no other indication if a VAX node fails.

WHEN THE VAXcluster SYSTEM HANGS

The following conditions may cause the cluster to hang:

- Cluster quorum is lost
- Shared system resource inaccessible
- Trying to boot from system disks with duplicate volume labels

labels moeten uniek zijn over het hele cluster

Cluster Quorum is Lost

- Process creation and I/O activity on all nodes ceases.
- Information about events that cause loss of quorum are sent to OPCOM and OPA0 at each node.
  - Quorum may be lost before OPCOM has a chance to display the message.
  - OPA0 is the most reliable source of information.
- You could lower quorum and reconfigure the cluster or you could have a node rejoin the cluster.

WARNING

- **Be patient. Heavy network traffic or a long transition may make the VAXcluster system appear stopped.**
- **Be sure you know why you are lowering quorum if you choose to do so. It is not something to do routinely.**

Shared Resource Inaccessible

- Access to shared resources is managed by the Distributed Lock Manager (DLU).
- When a process is granted a lock on a resource, other processes must wait to access it
- If the lock is held for an extended period of time, the cluster will appear to hang.

Conditions that may cause this to happen:

- Volume rebuilding causes the volume being rebuilt to be locked.
- Locks on parts of SYS$SYSTEM:SYSUAF.DAT for write access will cause processes attempting to log in to appear to hang.
  - Normally, locks such as this do not endure long enough to cause major problems.
  - If the process with the lock is unable to proceed, then the backlog of blocked processes could proliferate rapidly.
  - The lock manager is functioning correctly in this case, so no error messages are generated.

Booting with Duplicate System Disk Volume Labels

Volume labels must be unique in a VAXcluster system.

- Attempting to boot a system with a duplicate system disk volume label causes the system to hang.
- Learn to recognize the console messages.

A possible solution:

- Boot one system and use it to issue a SET VOLUME/LABEL on one of the disks with duplicate labels.

Example 7–5 Duplicate System Disk Volume Label Messages

```
waiting to form or join VAXcluster
%CNXMAN, Discovered system VAXB
%CNXMAN, Established connection to system VAXB
%CNXMAN, Sending VAXcluster membership request to system VAXB
%CNXMAN, Now a VAXcluster member -- system VAXA
%SYSINIT-E- error mounting system device, status = 007280B4
%SYSINIT-E- error opening or mapping F11BXQP, status = 00018272
%SYSINIT-E- message file not found, or insufficient SPT to map it, status
= 00018272
```

write sys$output f$message(%X0078~~0~~B4)

another volume of same label already mounted

VAXcluster SPECIFIC ERROR MESSAGES

VAXport Errors

- Error messages generated by the CI DSSI port are prefixed with *%PAA0.*
- Error messages generated by the Ethernet port are prefixed with *%PEA0.*
- These messages appear in the error log and at the console.
  - Use VAXsimPLUS or SHOW ERROR to see errors.
  - Use ANALYZE/ERROR_LOG to examine error log.
- Some messages also appear on the console terminal.
  - All errors that cannot be logged
  - Some errors that are always displayed
- Example of VAXport error message:

```
DATA CABLE(S) CHANGE OF STATE
PATH 0.  WENT FROM GOOD TO BAD
```

  - Explanation:

    The VAXport driver logs this event which indicates some type of hardware problem with the CI, cables, or ports.

  - User Action:

    Check path A cables to see that they are not broken or improperly connected.

  Another error message:

```
DATAGRAM FREE QUEUE INSERT FAILURE
```

  - Explanation:

    The VAXport driver attempts to reinitialize the port.
    After 50 failing attempts, it marks the device off-line.
    Possible sources of the problem are CI hardware failures or memory, SBI, VAX–11/780, CMI, VAX–11/750, or BI, VAX 8200, VAX 8300, and VAX 8800 contention.

  - User Action:

    Call Digital Customer Service.
    This error is caused by a failure to obtain access to an interlocked queue.

Error Log Entries for Cluster-Available Devices

HSC device errors are:

- Logged by the VAX system that requested the associated I/O operation.
- If not associated with I/O operation, logged as VAXport error.
  (HSC ERROR LOGGING DATAGRAM)

MSCP served disk errors are:

- Logged by the system on which the disk is local
- If dual-ported, logged by the system that served the associated I/O operation

NODE FAILURES

There are many reasons why a node can fail, but they are not all cluster-specific and are therefore not covered in this course.

CLUEXIT Bugchecks

The VMS operating system performs bugchecks when it detects conditions that could compromise the integrity of system activity.

Conditions that cause a bugcheck include:

- Configuration errors
- System management errors
- Hardware failures

Some of the most common conditions that result in CLUEXIT bugchecks include:

- Cluster connection between nodes is broken for a longer interval than RECNXINTERVAL seconds.
- Cluster partitioning occurs.
- The value of SYSGEN parameter SCSMAXMSG for a node is too small.

GATHERING DIAGNOSTIC INFORMATION

The following are methods of obtaining diagnostic information:

- SHOW DEVICE
- SHOW CLUSTER
- MONITOR
- Examining SYSGEN parameters
- System Dump Analyzer (SDA)
- The VAXsimPLUS utility (covered in the Appendix of this module)

Using SHOW DEVICE in a VAXcluster System

- SHOW DEVICE/FULL
  - Shows the complete status of a device
  - Useful for determining the configuration of disks in a cluster
- SHOW DEVICE/FILES
  - This command lists files opened only on this node.
  - To find all open files on a disk, use either the SHOW DEVICE/FILES command or SYSMAN commands on each node that has the disk mounted, or SYSMAN.
- SHOW DEVICE/SERVED
  - Shows information about MSCP served disks on the local system

SHOW DEVICE/FULL

Cluster-related disk device information:

- Whether the disk is available to the cluster
- Whether it is either MSCP served and/or dual-ported (local disks only)
- The name and type, VAX or HSC of primary and secondary hosts
- Whether the disk is mounted on this system
- The systems in the cluster on which the disk is mounted

Examples 7–6 through 7–8 illustrate output produced by SHOW DEVICE/FULL for various types of disks.

Example 7-6 SHOW DEVICE/FULL for MSCP Served Disk

```
$ SHOW DEVICE/FULL $1$DUA0

Disk $1$DUA0:(BARNUM), device type RA81, is online, mounted, file-oriented device,
    shareable, served to cluster via MSCP Server, error logging is enabled.

    Error count                    0    Operations completed               5989
    Owner process                 ""    Owner UIC                         [1,1]
    Owner process ID        00000000    Dev Prot    S:RWED,O:RWED,G:RWED,W:RWED
    Reference count                1    Default buffer size                 512
    Total blocks              891072    Sectors per track                    51
    Total cylinders             1248    Tracks per cylinder                   8
    Allocation class               1

    Volume label            "FLYING"    Relative volume number                0
    Cluster size                   3    Transaction count                    93
    Free blocks                 8069    Maximum files allowed            222768
    Extend quantity                5    Mount count                           7
    Mount status              System    Cache name           "_$1$DUA0:XQPCACHE"
    Extent cache size             64    Maximum blocks in extent cache     2088
    File ID cache size            64    Blocks currently in extent cache      0
    Quota cache size              30    Maximum buffers in FCP cache        129

  Volume status:  subject to mount verification, file high-water marking, write-
      through caching enabled.
  Volume is also mounted on RNGLNG, BAILEY, LION, HORSE, BEAR, TIGER.
```

Example 7-7 SHOW DEVICE/FULL for MSCP Served Disk Local to Another System

```
$ SHOW DEVICE/FULL $1$DUA2

Disk $1$DUA2:(BARNUM), device type RA81, is online, mounted, file-oriented
    device, shareable, available to cluster, error logging is enabled.

    Error count                   0    Operations completed             5989
    Owner process                ""    Owner UIC                       [1,1]
    Owner process ID       00000000    Dev Prot    S:RWED,O:RWED,G:RWED,W:RWED
    Reference count               1    Default buffer size               512
    Total blocks             891072    Sectors per track                  51
    Total cylinders            1248    Tracks per cylinder                 8
    Host name              "BARNUM"    Host type, available     VAX 8810, yes
    Allocation class              1

    Volume label           "FLYING"    Relative volume number              0
    Cluster size                  3    Transaction count                  93
    Free blocks                8069    Maximum files allowed          222768
    Extend quantity               5    Mount count                         7
    Mount status             System    Cache name          "_$1$DUA2:XQPCACHE"
    Extent cache size            64    Maximum blocks in extent cache   2088
    File ID cache size           64    Blocks currently in extent cache    0
    Quota cache size             30    Maximum buffers in FCP cache      129

  Volume status:  subject to mount verification, file high-water marking, write-
      through caching enabled.
  Volume is also mounted on BARNUM, RNGLNG, LION, HORSE, BEAR, TIGER.
```

Example 7–8 SHOW DEVICE/FULL for Dual-Ported HSC Disk

```
$ SHOW DFVICE/FULL $1$DUA1

Disk $1$DUA1: (HIWIRE), device type RA82, is online, mounted, file-oriented
    device, shareable, available to cluster, error logging is enabled.

    Error count                      2    Operations completed          2352587

    Owner process                   ""    Owner UIC                       [1,1]
    Owner process ID          00000000    Dev Prot     S:RWED,O:RWED,G:RWED,W:RWED
    Reference count                 76    Default buffer size               512
    Total blocks               1216665    Sectors per track                  51
    Total cylinders               1248    Tracks per cylinder                14
    Host name                 "HIWIRE"    Host type, available       HSC50, yes
    Alternate host name        "CLOWN"    Host type, available       HSC70, yes
    Allocation class                 1

    Volume label           "TIGHTROPE"    Relative volume number              0
    Cluster size                     1    Transaction count                 223
    Free blocks                 121857    Maximum files allowed          222768
    Extend quantity                  5    Mount count                         7
    Mount status                System    Cache name         "_$1$DUA1:XQPCACHE"
    Extent cache size               64    Maximum blocks in extent cache    806
    File ID cache size              64    Blocks currently in extent cache   70
    Quota cache size                 0    Maximum buffers in FCP cache      350

  Volume status:  subject to mount verification, file high-water marking, write-
     through caching enabled.
  Volume is also mounted on BAILEY, LION, HORSE, RNGLNG, BEAR, TIGER.
```

SHOW DEVICE/FILES

Shows which files on a particular device are being accessed by processes on the local system. Do this from SYSMAN to get a picture of all the activity.

Example 7–9 SHOW DEVICE/FILES Output

```
$ SHOW DEVICE/FILE $1$DUA1:

Files accessed on device _$1$DUA1: on  9-MAY-1989 11:47:44.50

Process name       PID      File name
                 00000000  [000000]INDEXF.SYS;1
                 00000000  [000000]QUOTA.SYS;1
T O M            20202580  [JAGGER.MAIL]MAIL.MAI;1
Bette            2020267C  [FINNERN.OLTP]OBJECTIVES.TJL;1
Ed Bernstein     2020233A  [BERNSTEIN]MYEVEPLUS.TPU$SECTION;44
Ed Bernstein     2020233A  [BERNSTEIN.CLUSTER]LL1.TJL;1
Mike Beeler      2020275B  [BEELER.CLUSTER.LL]TEST_IT.DAT;1
Mike Beeler      2020275B  [BEELER.CLUSTER.LL]D.COM;1
Mike Beeler      2020275B  [BEELER.CLUSTER.LL]T.LIS;1
```

SHOW DEVICE/SERVED

Displays information on devices served by the MSCP server on this node.

Some of the qualifiers you might use with SHOW DEVICE/SERVED:

- /ALL gives all information listed below.
- /HOST displays the names of processors that have devices on-line through the local MSCP server, and the number of devices.
- /RESOURCE displays the resources available to the MSCP server: total amount of nonpaged dynamic memory available for I/O buffers, number of I/O request packets.
- /COUNT displays the number of each size and type of I/O operation the MSCP server has performed since it was started.

MONITOR MSCP initiates monitoring of the I/O operations that use the MSCP Server.

Example 7-10 SHOW DEVICE/SERVED Output

```
$ SHOW DEVICE /SERVED /ALL

      MSCP-Served Devices on BARNUM  6-JUN-1989 16:21:06.22

                                              Queue Requests
Device:          Status      Total Size      Current    Max     Hosts
    DUA2         Online          891072            0      2         6
    DUA0         Online         1216665            0      2         6
    DUA1         Online         1216665            0      2         6

                                              Queue Requests
Host:              Time of Connection        Current    Max   Devices
    TIGER        6-JUN-1989 09:38:01.79            0      3         3
    HORSE        6-JUN-1989 11:10:14.02            0      3         3
    RNGLNG       6-JUN-1989 10:31:32.02            0      3         3
    BAILEY       6-JUN-1989 10:41:56.32            0      3         3
    LION         6-JUN-1989 09:45:34.95            0      3         3
    BEAR         6-JUN-1989 11:00:26.94            0      3         3

Resources:         Total        Free       In Use
    Buffer Area:     128         128            0
    I/O Packets:       0           0

                 Current     Maximum
    Buffer Wait:       0           1

Request Count:
       0-7:     34223         32-39:      218         88-103:      380
      8-15:      3100         40-55:      155        104-127:       66
     16-23:      3577         56-71:      154
     24-31:       239         72-87:       86

Operations Count:
     ABORT                0     ERASE             27     READ          39243
     ACCESS               0     FLUSH              0     REPLACE           0
     AVAILABLE            7     GET COM STS        0     SET CTL CHR      12
     CMP CTL DAT          0     GET UNT STS     2328     SET UNT CHR       5
     CMP HST DAT          0     ONLINE             9     WRITE          2955
   Total              44586
```

Using SHOW CLUSTER in a VAXcluster System

SHOW CLUSTER displays a variety of information about the cluster. It provides a view of the cluster as seen from a single node, rather than a complete view of the cluster.

There are two types of displays:

- One time display (SHOW CLUSTER)
- Dynamic display (SHOW CLUSTER/CONTINUOUS)

The SHOW CLUSTER display can be modified to include any desired information.

Example 7–11 The Default SHOW CLUSTER Display

```
$ SHOW CLUSTER
View of Cluster from system ID 1025   node: BARNUM     12-MAY-1990 17:09:35

        +-----------------------------+
classes |       SYSTEMS      | MEMBERS |
        +-----------------------------+
fields  | NODE   | SOFTWARE |  STATUS |
        +-----------------------------+
        | BARNUM | VMS V5.4 | MEMBER  |
        | CLOWN  | HSC V390 |         |
        | BAILEY | VMS V5.4 | MEMBER  |
        | HIWIRE | HSC V390 |         |
        +-----------------------------+
```

BRK_NON is van een machine die uit het cluster is gegaan.

- SYSTEMS and MEMBERS are classes of information.
  - — NODE and SOFTWARE are fields within the SYSTEMS class.
  - — STATUS is a field within the MEMBERS class.

You can control the SHOW CLUSTER display with the commands listed in Table 7–1.

Table 7–1 Basic SHOW CLUSTER Commands

Command	Description
ADD	Add a class or field to the display
REMOVE	Remove a class or field from the display
SET	Change the width or characteristics of a field
INIT	Reset the display to a known
HELP	Enter interactive help mode
EXIT	Exit the display
SAVE	Create a command procedure SHOW_CLUSTER.COM, which recreates the state of your screen
WRITE	Write current data to a file for problem reports
@file-spec	Execute a command procedure of SHOW CLUSTER commands

Logical: show_cluster $init dub3:[Gortemaker.test]
Show_cluster.com

The SHOW CLUSTER utility can display several classes of information, as shown in Table 7–2.

Table 7–2 Classes of Data Displayed by the SHOW CLUSTER Utility

Class	Description
CLUSTER	General information about the cluster
SYSTEMS	Known systems in a cluster
MEMBERS	Systems actively participating in the cluster
CIRCUITS	Communication paths between systems
CONNECTIONS	Connections established over a communication path
COUNTERS	Traffic counts for each connection
CREDITS	Credit counts for each connection
LOCAL_PORT	VAXport hardware on the local system
ERROR	Error status for the local VAXport

Example 7–12 SHOW CLUSTER Output (MEMBERS)

```
Command> INIT
Command> REMOVE SOFTWARE,STATUS
Command> ADD VOTES,EXPECTED_VOTES,QF_SAME,QF_ACTIVE
View of Cluster from system ID 1025  node: BARNUM  12-MAY-1990 15:31:31
+--------+-------------------------------------+
| SYSTEMS|              MEMBERS                |
+--------+-------+--------+---------+----------+
|  NODE  | VOTES | EXPECT | QF_SAME | QF_ACTIVE |
+--------+-------+--------+---------+----------+
| BARNUM |     1 |      2 | YES     | YES      |
| BAILEY |     1 |      2 | YES     | YES      |
| CLOWN  |       |        |         |          |
| HIWIRE |       |        |         |          |
+--------+-------+--------+---------+----------+
```

Example 7–13 SHOW CLUSTER Output (Other MEMBERS Fields)

```
Command> REMOVE MEMBERS
Command> ADD CSID,CNX_STATE,TRANSITION_TIME

View of Cluster from system ID 1025  node: BARNUM   12-MAY-1990 15:31:13
+--------+------------------------------------+
| SYSTEMS|            MEMBERS                 |
+--------+--------+--------+-----------------+
|  NODE  |  CSID  | CNX_ST | TRANSITION_TIME |
+--------+--------+--------+-----------------+
| BARNUM | 10009  | NEW    | 29-APR-90 21:19 |
| BAILEY | 1000A  | OPEN   |  3-MAY-90 08:56 |
| CLOWN  |        |        |                 |
| HIWIRE |        |        |                 |
+--------+--------+--------+-----------------+
```

Example 7–14 SHOW CLUSTER Output (CIRCUITS)

```
$ SHOW CLUSTER/CONTINUOUS
Command> INIT
Command> ADD CIRCUITS

View of Cluster from system ID 1025  node: BARNUM      12-MAY-1990 17:19:16

+-------------------+---------+----------------------+
|      SYSTEMS      | MEMBERS |       CIRCUITS       |
+--------+----------+---------+------+-------+-------+
|  NODE  | SOFTWARE |  STATUS |RPORT |RP_TYP |CIR_STA |
+--------+----------+---------+------+-------+-------+
| BARNUM | VMS V5.4 | MEMBER  |    1 |CI780  | OPEN   |
| CLOWN  | HSC V390 |         |    3 |HSC50  | OPEN   |
| BAILEY | VMS V5.4 | MEMBER  |    2 |CI780  | OPEN   |
| TIGER  | VMS V5.4 | MEMBER  |      |ETHERN | OPEN   |
| HIWIRE | HSC V390 |         |    4 |HSC50  | OPEN   |
+--------+----------+---------+------+-------+-------+
```

Example 7–15 SHOW CLUSTER Output (CONNECTIONS)

```
Command> REMOVE MEMBERS,CIRCUITS,SOFTWARE
Command> ADD CONNECTIONS,REM_PROC_NAME

View of Cluster from system ID 1025  node: BARNUM      12-MAY-1990 09:53:08
+---------+-------------------------------------------------+
| SYSTEMS |                   CONNECTIONS                   |
+---------+------------------+-------------------+----------+
|  NODE   |  LOC_PROC_NAME   |   REM_PROC_NAME   | CON_STA  |
+---------+------------------+-------------------+----------+
| BARNUM  | SCS$DIRECTORY    |         ?         | LISTEN   |
|         | MSCP$TAPE        |         ?         | LISTEN   |
|         | VMS$VAXcluster   |         ?         | LISTEN   |
|         | SCA$TRANSPORT    |         ?         | LISTEN   |
|         | MSCP$DISK        |         ?         | LISTEN
| CLOWN   | VMS$DISK_CL_DRVR | MSCP$DISK         | OPEN     |
| BAILEY  | VMS$DISK_CL_DRVR | MSCP$DISK         | OPEN     |
|         | MSCP$DISK        | VMS$DISK_CL_DRVR  | OPEN     |
|         | VMS$VAXcluster   | VMS$VAXcluster    | OPEN     |
| TIGER   | MSCP$DISK        | VMS$DISK_CL_DRVR  | OPEN     |
|         | VMS$VAXcluster   | VMS$VAXcluster    | OPEN     |
| HIWIRE  | VMS$DISK_CL_DRVR | MSCP$DISK         | OPEN     |
+---------+------------------+-------------------+----------+
```

SHOW CLUSTER Initialization Files

The SHOW CLUSTER utility enables you to specify an initialization file to be used when the utility is invoked.

- Define the logical name SHOW_CLUSTER$INIT to point to this file.
- The sample initialization file shown in Example 7–16 illustrates the command format, and Example 7–17 shows the display that is defined by the sample.

Example 7–16 A Sample SHOW CLUSTER Initialization File

```
INITIALIZE
ADD CL_EXPECTED_VOTES,CL_QUORUM,CL_VOTES,CL_QDVOTES,CL_MEMBERS
ADD LAST_TRANSITION,CNX_STATE,VOTES,EXPECTED_VOTES,CIR_STATE,CABLE_STATUS
SET CL_QUORUM /WIDTH =   4
SET CL_VOTES /WIDTH =   4
SET VOTES /WIDTH =   1
SET EXPECTED_VOTES /WIDTH =   1
SET CIR_STATE /WIDTH =   5
```

Example 7–17 Display Resulting from SHOW_CLUSTER$INIT

```
View of Cluster from system ID 19750  node: BARNUM    6-MAY-1990 15:42:35
+------------------+-------------------------+---------------+
|      SYSTEMS     |         MEMBERS         |    CIRCUITS   |
+--------+---------+--------+---+---+--------+-------+-------+
| NODE   | SOFTWARE| CNX_ST | V | E | STATUS | CIR_S | CABLE |
+--------+---------+--------+---+---+--------+-------+-------+
| BARNUM | VMS V5.4| NEW    | 1 | 3 | MEMBER | OPEN  | A - B |
| HIWIRE | HSC V390|        |   |   |        | OPEN  | A - B |
| TIGER  | VMS V5.4| OPEN   | 0 | 3 | MEMBER | OPEN  |       |
| HORSE  | VMS V5.4| OPEN   | 0 | 3 | MEMBER | OPEN  |       |
| BAILEY | VMS V5.4| OPEN   | 1 | 3 | MEMBER | OPEN  | A - B |
| CLOWN  | HSC V390|        |   |   |        | OPEN  | A - B |
+--------+---------+--------+---+---+--------+-------+-------+

+------------------------------------------------------------+
|                          CLUSTER                           |
+--------+------+------+--------+------------+---------------+
| CL_EXP | CL_Q | CL_V | CL_QDV | CL_MEMBERS | LAST_TRANSITION |
+--------+------+------+--------+------------+---------------+
|      3 |    2 |    3 |  65535 |          4 | 6-MAY-90 11:10 |
+--------+------+------+--------+------------+---------------+
```

Using MONITOR in a VAXcluster System

MONITOR classes that are cluster-specific:\bold)

- CLUSTER gives statistics over the entire cluster.
- MSCP gives MSCP server statistics.
- SCS gives system communication services statistics.
- DLOCK gives distributed lock management statistics.

MONITOR qualifiers that are useful in a cluster:

- /NODE gives statistics from particular nodes.
- /BY_NODE gives a multinode summary.

MONITOR CLUSTER Command

Example 7-18 MONITOR CLUSTER Output

hiervoor is DECNET nodig.

```
$ MONITOR
MONITOR> MONITOR CLUSTER
%MONITOR-I-ESTABCON, establishing connection to remote nodes...

Statistic: CURRENT            VAX/VMS Monitor Utility        9-JUN-1989 16:55:47
                                  CLUSTER STATISTICS
                   CPU                    |                   MEMORY
                                          |
CPU Busy             0   25   50   75  100|%Memory In Use   0   25   50   75  100
                     +----+----+----+----+|                 +----+----+----+----+
HORSE           100 |********************|HORSE         94 |******************
TIGER            19 |***                 |TIGER         92 |******************
BARNUM           11 |**                  |LION          64 |************
BAILEY            5 |*                   |BAILEY        43 |********
RNGLNG            4 |                    |RNGLNG        32 |******
LION              4 |                    |BARNUM        24 |****
BEAR              4 |                    |BEAR          22 |****
------------------------------------------+-------------------------------------
                   DISK                   |                    LOCK
                                          |
I/O Operation Rate   0   25   50   75  100|Tot ENQ/DEQ Rate 0  125  250  375  500
                     +----+----+----+----+|                 +----+----+----+----+
$1$DUA0:          6 |*                   |BARNUM         8 |
$1$DUA0:     R    6 |*                   |RNGLNG           |
$1$DUA1:          2 |                    |TIGER            |
$1$DUA1:     R    1 |                    |HORSE            |
$1$DUA2:            |                    |LION             |
$1$DUA3:            |                    |BAILEY           |
                                          |BEAR             |
```

MONITOR MSCP Command

Example 7–19 MONITOR MSCP Output

```
                          VAX/VMS Monitor Utility
                           MSCP SERVER STATISTICS
                              on node WHYNOT
                           12-JUN-1989 13:52:03
                                    CUR        AVE        MIN        MAX

Server I/O Request Rate             7.93       2.50       0.00       7.93
Read Request Rate                   7.93       2.50       0.00       7.93
Write Request Rate                  0.00       0.00       0.00       0.00

Extra Fragment Rate                 0.00       0.00       0.00       0.00
Fragmented Request Rate             0.00       0.00       0.00       0.00
Buffer Wait Rate                    0.00       0.00       0.00       0.00

Request Size Rates         1        7.93       2.18       0.00       7.93
 (Blocks)                2-3        0.00       0.12       0.00       0.32
                         4-7        0.00       0.06       0.00       0.32
                        8-15        0.00       0.12       0.00       0.32
                       16-31        0.00       0.00       0.00       0.00
                       32-63        0.00       0.00       0.00       0.00
                         64+        0.00       0.00       0.00       0.00
```

MONITOR Multinode Summary

MONITOR gathers data on only one node.

- You can generate a multinode summary for any MONITOR class. This is especially useful for monitoring total disk activity.

For a side-by-side display of MONITOR data from multiple nodes:

- Use MONITOR to collect data in a file on each node.
- Then use MONITOR/SUMMARY/BY_NODE to combine the data files and produce a multinode summary.

MONITOR /NODE Qualifier

Example 7–20 Using the /NODE Qualifier with MONITOR DISK

```
MONITOR> MONITOR/NODE=(BARNUM,BAILEY,RNGLNG) DISK
%MONITOR-I-ESTABCON, establishing connection to remote nodes...
                        VAX/VMS Monitor Utility
                    DISK I/O STATISTICS on node BARNUM
                          20-DEC-1989 11:25:11

I/O Operation Rate                            CUR        AVE        MIN        MAX

$1$DUA0:      (CLOWN)  BARNUM_SYS            0.59       6.90       0.00      16.75
$1$DUA1:      (CLOWN)  TIGHTROPE             0.00       0.15       0.00       0.94
$1$DUA2:      (BARNUM) FLYING                0.00       0.00       0.00       0.00
$1$DUA3:      (BAILEY) TRAPEZE               2.09       3.31       0.00      11.59
$2$DUA0:      (HORSE)  THREE                 0.00       0.03       0.00       1.21
$2$DUA1:      (HORSE)  RING                  0.00       0.00       0.00       0.00

                        VAX/VMS Monitor Utility
                    DISK I/O STATISTICS on node BAILEY
                          20-DEC-1989 11:25:18

I/O Operation Rate                            CUR        AVE        MIN        MAX

$1$DUA0:      (CLOWN)  BARNUM_SYS            0.00       0.16       0.00       1.89
$1$DUA1:      (CLOWN)  TIGHTROPE             0.00       0.05       0.00       0.63
$1$DUA2:      (BARNUM) FLYING                0.00       0.00       0.00       0.00
$1$DUA3:      (BAILEY) TRAPEZE               0.00       0.00       0.00       0.00
$2$DUA0:      (HORSE)  THREE                 0.00       0.00       0.00       0.00
$2$DUA1:      (HORSE)  RING                  0.00       0.00       0.00       0.00

                        VAX/VMS Monitor Utility
                    DISK I/O STATISTICS on node RNGLNG
                          20-DEC-1989 11:25:16

I/O Operation Rate                            CUR        AVE        MIN        MAX

$1$DUA0:      (CLOWN)  BARNUM_SYS            2.18       1.82       0.00       9.09
$1$DUA1:      (CLOWN)  TIGHTROPE             0.31       0.02       0.00       0.31
$1$DUA2:      (BARNUM) FLYING                0.00       0.63       0.00       7.83
$1$DUA3:      (BAILEY) TRAPEZE               5.00       1.39       0.00       7.82
$2$DUA0:      (HORSE)  THREE                 0.00       0.00       0.00       0.00
$2$DUA1:      (HORSE)  RING                  0.00       0.00       0.00       0.00
```

Using MONITOR to Locate Disk I/O Bottlenecks

I/O bottlenecks can cause the system to appear to hang.

1. Determine which cluster-wide disks may be problem disks.

   - Create a node-by-node summary of disk I/O using MONITOR/NODE
   - Note disks with the row sum more than eight I/Os per second.
   - Eliminate from the list of cluster problem disks those that are:
     - Not cluster-accessible
     - Dedicated to an application
     - Being backed up

2. For each node, determine the impact of potential problem disks:

   - If a disproportionate amount of a disk's I/O comes from a particular node, the problem is probably specific to the node.
   - If a disk's I/O is spread evenly over the cluster, the problem may be cluster-wide overuse.
   - If the average queue length for a disk on a given node is less than 0.2, then the disk is having little impact on the node.

3. For each problem disk, determine whether:

   - Page and swap files from any node are on the disk.
   - Multiple operating systems are on the disk.
   - Commonly used programs or data files are on the disk (SHOW DEVICE/FILES).
   - Users with default directories on the disk are causing the problem.

Example 7-21 Command Procedure for MONITOR Recording

```
$! Get the node name to use in the recording file name
$!
$ NODE :== F$GETSYI("NODENAME")
$!
$! Record disk I/O rates for the next hour
$!
$ MONITOR /NODISPLAY /END="+1:00" /RECORD='NODE'.MON DISK
```

Example 7-22 Procedure to Start Recording on Two Nodes

```
$! Record data in batch on node BARNUM
$!
$ SUBMIT /QUEUE=BARNUM_BATCH RECORD.COM
$!
$! Record data in batch on node BAILEY
$!
$ SUBMIT /QUEUE=BAILEY_BATCH RECORD.COM
```

Example 7-23 Two-Node MONITOR Summary

```
$ MONITOR /SUMMARY /BY_NODE /INPUT=(BARNUM.MON,BAILEY.MON) /NODISPLAY DISK
$ TYPE MONITOR.SUM
-------               VMS Monitor Utility
| AVE |               DISK I/O STATISTICS
-------                MULTI-FILE SUMMARY
I/O Operation Rate

Node:       BARNUM              BAILEY
From: 3-DEC-1989 15:06  3-DEC-1989 15:06  Row    Row      Row      Row
To:   3-DEC-1989 16:06  3-DEC-1989 16:06  Sum  Average  Minimum  Maximum

$1$DUA0:              0.34              0.00  0.3      0.1     0.00     0.34
$1$DUA1:              0.02              1.01  1.0      0.5     0.02     1.01
$1$DUA2:              1.13              3.48  4.6      2.3     1.13     3.48
$1$DUA3:              8.24             12.42 20.6     10.3     8.24    12.42
$2$DUA0:              0.00              0.00  0.0      0.0     0.00     0.00
$2$DUA1:              0.00              0.00  0.0      0.0     0.00     0.00
```

NOTE

This example does not show the entire display because it would require 132 columns.

Examining SYSGEN Parameters

There are three ways to display and modify VAXcluster related parameters.

- SYSBOOT for problems that occur during conversational boot
- SYSGEN for local parameters
- SYSMAN for cluster-wide parameters

The two parameter classes of interest are:

- /CLUSTER
- /SCS

Example 7–24 illustrates the use of the SYSMAN utility to display the values of these parameters.

Example 7–24 Output from the PARAMETER SHOW/CLUSTER Command

```
$ RUN SYS$SYSTEM:SYSMAN
SYSMAN> PARAMETERS SHOW /CLUSTER
%SYSMAN-I-USEACTNOD, a USE ACTIVE has been defaulted on node HORSE
Node HORSE:   Parameters in use: ACTIVE
Parameter Name             Current   Default   Minimum   Maximum Unit  Dynamic
--------------             -------   -------   -------   ------- ----  -------
❶ VAXCLUSTER                     2         1         0         2 Coded-value
❷ EXPECTED_VOTES                 3         1         1       127 Votes
VOTES                            0         1         0       127 Votes
RECNXINTERVAL                   20        20         1     32767 Seconds      D
DISK_QUORUM     "                "    "    "    "    "    "ZZZZ" Ascii
QDSKVOTES                        1         1         0       127 Votes
QDSKINTERVAL                    10        10         1     32767 Seconds
ALLOCLASS                        2         0         0       255 Pure-number
LOCKDIRWT                        0         0         0       255 Pure-number
NISCS_CONV_BOOT                  0         0         0         1 Boolean
❸ NISCS_LOAD_PEA0                1         0         0         1 Boolean
NISCS_PORT_SERV                  0         0         0         3 Bit-encoded
MSCP_LOAD                        1         0         0         1 Boolean
MSCP_SERVE_ALL                   2         0         0         2 Coded-value
MSCP_BUFFER                    128       128        16        -1 Coded-value
MSCP_CREDITS                     4         4         2         8 Coded-value
SYSMAN> PARAMETERS SHOW /SCS
%SYSMAN-I-USEACTNOD, a USE ACTIVE has been defaulted on node HORSE
Node HORSE:   Parameters in use: ACTIVE
Parameter Name             Current   Default   Minimum   Maximum Unit  Dynamic
--------------             -------   -------   -------   ------- ----  -------
SCSBUFFCNT                     512        50         0     32767 Entries
SCSCONNCNT                      40        40         2     32767 Entries
SCSRESPCNT                     300       300         0     32767 Entries
SCSMAXDG                       576       576        28       985 Bytes
SCSMAXMSG                      112       112        52       985 Bytes
SCSFLOWCUSH                      1         1         0        16 Credits      D
❹ SCSSYSTEMID                 1120         0        -1        -1 Pure-number
SCSSYSTEMIDH                     0         0        -1        -1 Pure-number
❺ SCSNODE                 "HORSE    "    "    "    "    "    "ZZZZ" Ascii
PRCPOLINTERVAL                  30        30         1     32767 Seconds      D
PASTIMOUT                        5         5         1        99 Seconds      D
PASTDGBUF                       16         4         1        16 Buffers
PANUMPOLL                       16        16         1       223 Ports        D
PAMAXPORT                       15        15         0       223 Port-number  D
PAPOLLINTERVAL                   5         5         1     32767 Seconds      D
PAPOOLINTERVAL                  15        15         1     32767 Seconds      D
PASANITY                         1         1         0         1 Boolean      D
PANOPOLL                         0         0         0         1 Boolean      D
UDABURSTRATE                     0         0         0        31 Longwords
```

Notes on Example 7–24:

❶ The VAXCLUSTER parameter must be set to a value of 1 or 2.

❷ The EXPECTED_VOTES parameter must be set to the maximum number of votes expected in the VAXcluster system.

❸ The NISCS_LOAD_PEA0 parameter must be set to 1 for all cluster nodes that establish an Ethernet connection to the cluster. Nodes that are booting from an HSC disk should set this parameter to 0.

❹ The SCSSYSTEMID parameter is an integer value that is also the DECnet node ID. This parameter is required for a node to join a cluster.

❺ The SCSNODE parameter must be set to the same value as the node name parameter provided by the NCP utility.

The System Dump File

- The system dump file contains a picture of what the system looks like when the cluster crashes.
- A dump of physical memory contents is performed to the file SYS$SYSTEM:SYSDUMP.DMP when the system fails abnormally and the SYSGEN parameter DUMPBUG is set to 1.
- The size of the dump files can limit the number of nodes that can boot from a single system disk.
- To allow more systems to share the same system disk and still have useful dump files that are smaller than all of system memory, use the SYSGEN parameter DUMPSTYLE.

For more information, refer to the *VMS System Dump Analyzer Utility Manual*.

Using the System Dump Analyzer in a VAXcluster System

The system dump analyzer can be used in one of two ways:

- To examine information about a system crash
- To examine elements of the system executive in a running VMS environment

Table 7–3 summarizes the SDA commands that relate to clusters.

Table 7–3 System Dump Analyzer Commands

Command	Information Displayed
SHOW CLUSTER	Cluster data structures on the local node
SHOW CONNECTIONS	Information about connections between SCS processes
SHOW LOCK	Information about resource locks
SHOW PORTS	Information about SCS ports
SHOW RESOURCES	Information resources
SHOW RSPID	Information about response-ids

To understand the information returned by SDA requires a knowledge of VMS operating system internals that is beyond the scope of this course.

TROUBLESHOOTING SUMMARY

If the source of a problem is not obvious:

- Get a "picture" of the VAXcluster configuration from the point of view of each node.
  - Run SHOW CLUSTER on each VAX node.
    Use SHOW CLUSTER/CONTINUOUS to monitor cluster formation.
  - Use the SHOW VIRTUAL command on each HSC unit.
  - Examine console output, VAX and HSC from each node.
  - Examine the operator log if a console listing is not available.
  - Examine the device configuration from each node:

    $ SHOW DEVICE PA, SHOW DEVICE PE

    $ SHOW DEVICE/FULL for disks

    HSC> SHOW DISK or SHOW TAPE

- Make sure SYSGEN and other parameters are correct:
  - Check SCSNODE, SCSSYSTEMID, ALLOCLASS.
  - Check HSC parameters: NAME, ID, ALLOCATE.
  - Most other cluster-related parameters should be the same on all nodes.
  - Check VAXCLUSTER, EXPECTED_VOTES, DISK_QUORUM.
  - Make sure there are unique volume labels for multiple system disks.
  - Examine the contents of registers in boot command procedures.
  - When booting a system with unknown SYSGEN parameters, perform a conversational boot to examine parameters.

SUMMARY

The single most important thing to know about your cluster when a problem occurs is what is normal for the cluster. If you know what is normal, you will more easily recognize what is abnormal.

When a problem occurs, examine the following diagnostic information:

- %CNXMAN messages
- VAXport logged error messages
- Other logged device error messages
- HSC error messages and fault codes

Use the following utilities to become familiar with the cluster and to help locate problems when they occur:

- HSC SETSHO
- SHOW DEVICE
- SHOW CLUSTER
- MONITOR
- System Dump Analyzer

To solve a VAXcluster problem:

- Wherever possible, effect your own solution using the tools described.
- Call the appropriate Digital representative.

FURTHER READINGS

- Become familiar with the VAXcluster Management section of the *VMS Version 5.4 Release Notes*, that contains hints not documented elsewhere.
- For a complete list of %CNXMAN messages, see the Appendix, Connection Manager Messages, of the *VMS VAXcluster Manual*.
- For help with CI troubleshooting, see the Appendix, VAXcluster Troubleshooting Information, of the *VMS VAXcluster Manual*.
- For complete documentation of the SHOW CLUSTER utility, see the *VMS Show Cluster Utility Manual*.
- See the *VMS DCL Dictionary* for a description of the SHOW DEVICE command.
- For complete documentation of HSC utility programs, read the *HSC User Guide*.
- If you wish to examine error log entries, refer to the *VMS Error Log Utility Manual*, which documents ANALYZE/ERROR_LOG.
- If you wish to examine system dumps or the internals of a running system, see the *VMS System Dump Analyzer Utility Manual*, which documents ANALYZE/CRASH_DUMP and ANALYZE/SYSTEM.
- Read the *VAXsimPLUS User Guide* for complete information on running the VAXsimPLUS utility, interpreting the output, and maintaining its database.
- For complete documentation of the MONITOR utility, refer to the *VMS Monitor Utility Manual*.

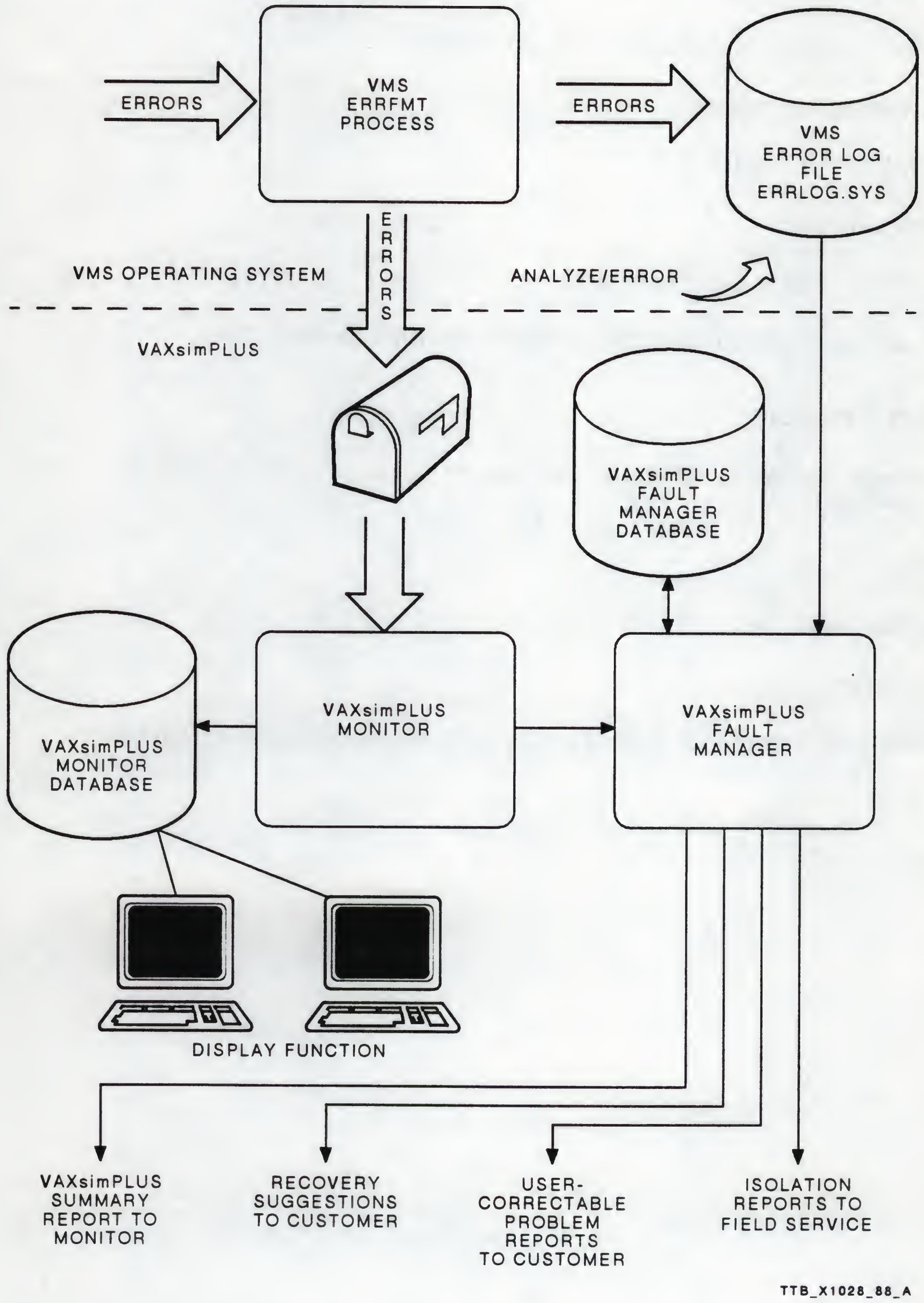
APPENDIX — USING THE VAXsimPLUS UTILITY in a VAXcluster SYSTEM

VAXsimPLUS (VAX System Integrity Monitor) software is a layered product that provides a graphic display of hardware status for the entire system or VAXcluster environment.

Features of VAXsimPLUS Software

- Monitors errors as they are logged
- Provides cursory rather than in-depth analysis
  - Does not replace ANALYZE/ERROR_LOG
- Not a cluster-specific utility
  - Especially useful in a cluster because it reduces a large amount of data to a simple display.
- The *VAXsimPLUS User Guide* describes how to:
  - Operate the display program
  - Maintain the VAXsimPLUS database

Figure 7–1 The Relationship Between VAXsimPLUS and VMS Software



VAXsimPLUS Data Collection

- The ERRFMT mailbox contains the current error log record.
- VAXsim_MONITOR detached process:
  - Attaches to the ERRFMT mailbox
  - Reads each mailbox entry
  - Filters out extraneous error log information
  - Creates and maintains its own historical database file (VAXsimDAT.DAT)

VAXsimPLUS Display Program

- Creates and maintains a display database in memory, using single or multiple VAXsimDAT.DAT files for:
  - Local nodes
  - Other VAXcluster nodes
  - Other DECnet nodes
- VAXsimPLUS software displays the following error information graphically in real time:
  - Error rates
  - Error log information

Figure 7–2 illustrates the information hierarchy for each system that VAXsimPLUS software is able to connect to.

Figure 7–2 VAXsimPLUS Information Hierarchy

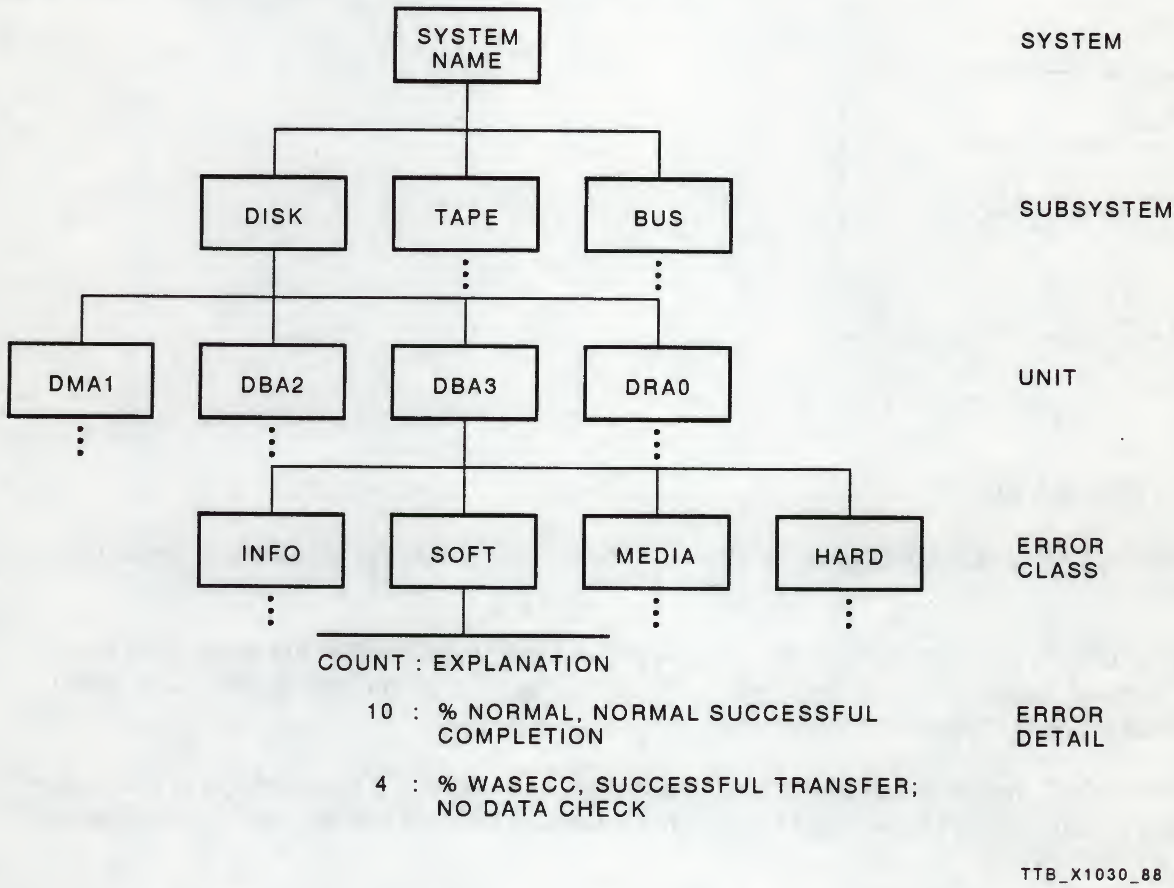
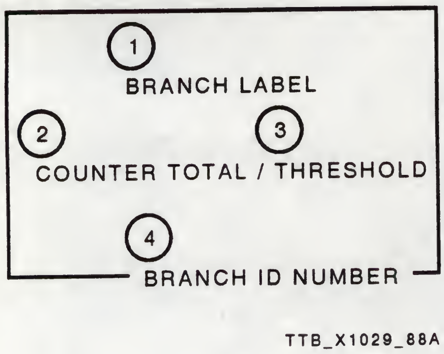


Figure 7–3 details the information elements for each item in the hierarchy.

Figure 7–3 VAXsimPLUS Block Diagram



Notes on Figure 7–3:

1. BRANCH LABEL identifies your location in the tree using names of nodes, subsystems, and devices.
2. COUNTER TOTAL summarizes the error activity of this branch within the evaluation period. The default period covers the previous 24 hours. Each counter total is the sum of all counter totals from the next lower level.
3. THRESHOLD appears only at the error classification level. It is a combination of the device average error rate and weighted margin and indicates how far the device is from triggering an error condition.
4. BRANCH ID NUMBER enables movement within the tree (when used with other commands).

Figure 7–4 Sample VAXsimPLUS Systems Display

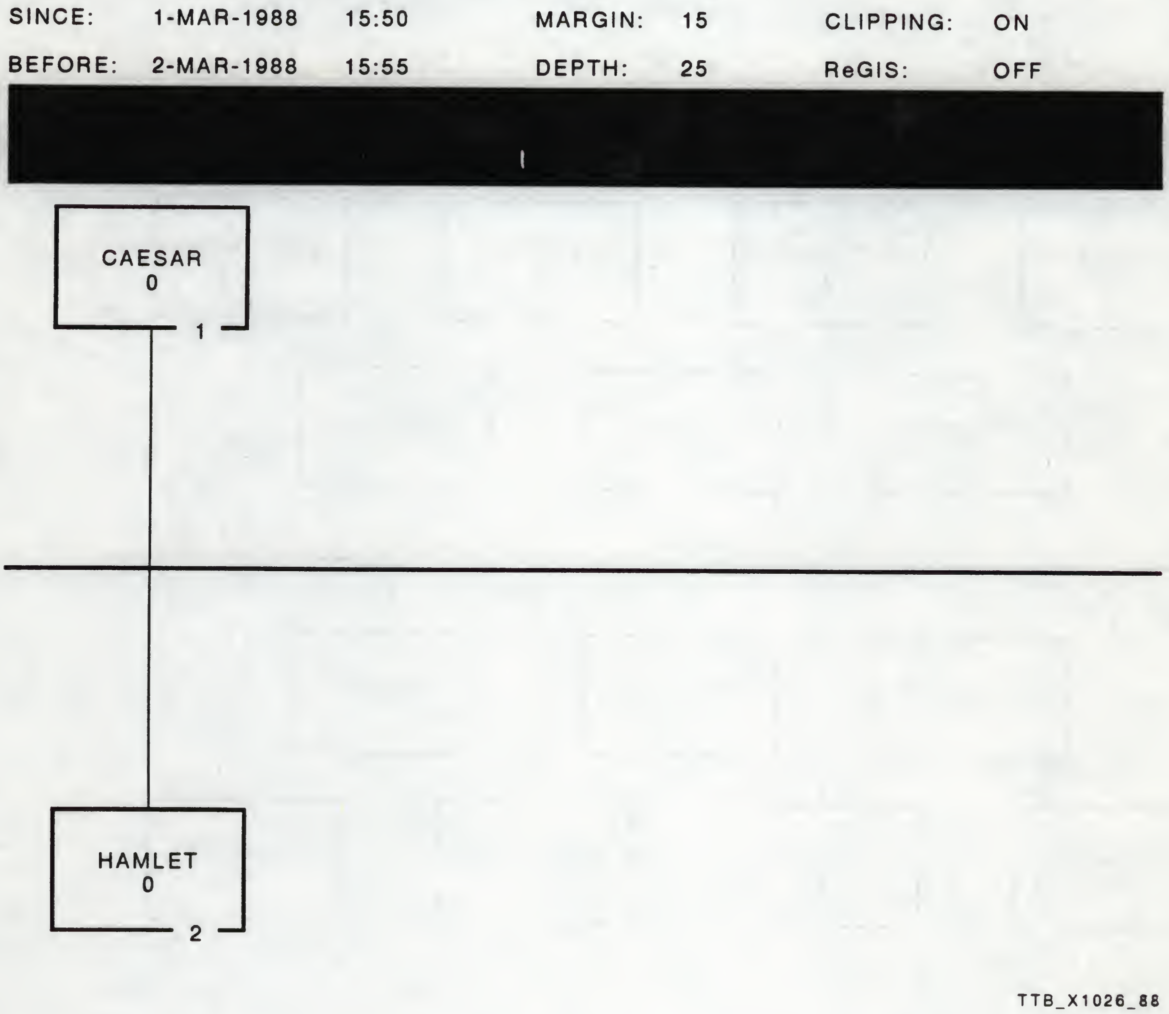


Figure 7-5 Sample VAXsimPLUS Disk Display

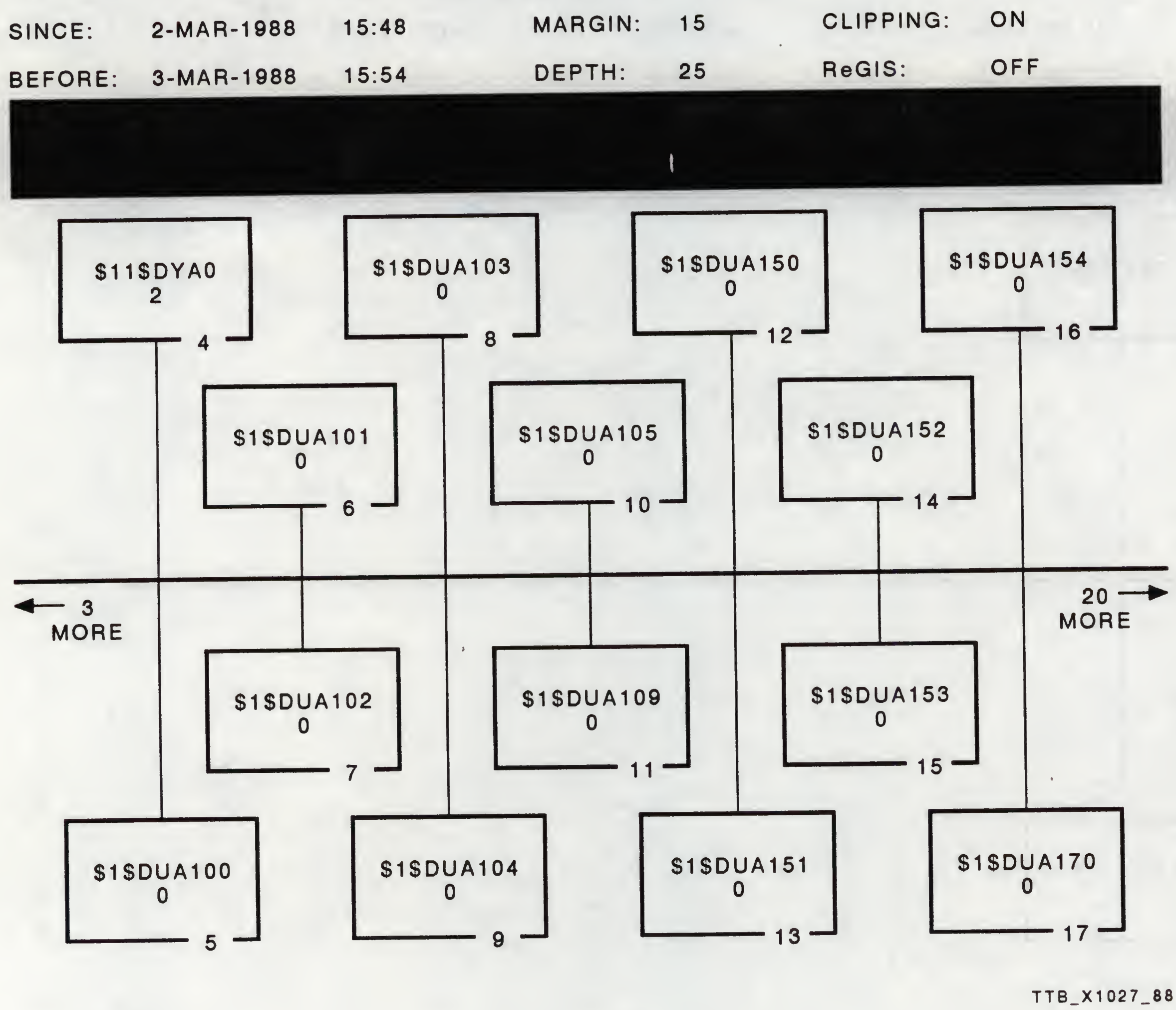


Table 7–4 summarizes the basic commands you can use to control the VAXsimPLUS display.

Table 7–4 Basic VAXsimPLUS Commands

Command	Description
CLUSTER	Displays general information about the VAXcluster system.
ADD	Adds information about other nodes to display.
REMOVE	Deletes information about other nodes from the display.
DOWN	Moves display position down to a more detailed level of information.
UP	Moves display position up by one or more display levels.
TOP	Moves display position to the topmost display.
ZOOM	Pinpoints an error condition by moving display position down as far as possible. ZOOM takes the single most obvious path if there is one; otherwise, it stops where more than one error status condition exists.
HELP	Provides help on VAXsimPLUS.
BEFORE	Sets ending time of evaluation period. Default BEFORE time is the current time.
SINCE	Sets beginning of evaluation period. Default SINCE time is approximately 24 hours earlier than BEFORE time.
UPDATE	Forces a reload of the entire database.
CYCLE	Cycles through all displays at the current level, updating the active database approximately every seven minutes.
WATCH	Monitors a single display, updating the active database approximately every seven minutes.

Exercises

PLANNING A VAXcluster SYSTEM

Written Exercise 1

Try answering the following questions about your own VAXcluster configuration. If you do not know enough about your own cluster to answer the questions, use the installation described below to answer them. Make your own assumptions about the installation if necessary.

1. Will you form a common-environment or mixed-environment cluster?

2. If you choose to form a mixed-environment VAXcluster system:

   — What resources (for instance: disk volumes, files) will be shared among nodes, if any?

   — Which users will be allowed access to which nodes, and why?

3. Which disks will be system disks? For which systems?

4. Would terminal servers be useful in your environment?

5. What performance bottleneck(s) might occur? How would you relieve them?

6. What additional hardware would you budget for?

Some of these questions may have more than one answer. It is recommended that you discuss the answers in class, especially if yours are different from the suggested answers.

Table 8–1 Characteristics of a Sample VAXcluster Configuration

	VAX 8820	VAX 8350	VAX 8350	HSC50
Node name	AL	DORA	BOB	JR
RP07	AL$DRA0:	-	-	-
RP07	AL$DRA1:	-	-	-
RA81	-	-	-	JR$DUA0:
RA81	-	-	-	JR$DUA1:
RA60	-	-	-	JR$DJA2:
TU78	AL$MFA0:	-	-	

University time-sharing environment:

- Student development of small programs
- Some administrative functions using database management software
- Communication by MAIL and PHONE

Node AL has been used for these functions for several years. DORA, BOB, and the HSC50 are new equipment.

Solutions to Exercise 1

Suggested answers for the sample installation:

1. In general, form a common-environment cluster unless you have a specific reason for forming a mixed-environment cluster. In the sample case, you can use the single existing user environment as a basis for a common-environment VAXcluster environment.

2. In the sample case, these questions do not apply.

3. A single common system disk is probably a good choice for the sample installation.

4. Terminal servers would be useful in the sample installation. The value of the common-environment cluster is that a user can work on any node; the terminal server makes it easy for a user to get to any node.

5. In the sample installation, almost any bottleneck could occur, depending on what the students are doing. The solution will vary depending on the bottleneck. Review the performance recommendations in Module 2 if necessary.

6. Some things that would be useful:

   - Dual-port kits for the RP07s, to provide two paths to each disk
   - TA78 upgrade kit for the TU78, so you can move it to the HSC50 and make it cluster-available
   - Another HSC controller, for redundancy
   - Additional VAX systems if you expect a CPU bottleneck
   - Additional disk drives if you expect an I/O bottleneck
   - Additional tape drives if you expect to make heavy use of tapes

Written Exercise 2

Interactive users can log in to VAXA and VAXB and also run batch jobs on them. VAXC is used for batch processing only.

Assign votes to each VAX node, and to a quorum disk if necessary, so that losing VAXC makes the VAXcluster no more likely to lose quorum than if VAXC were present. Also, tell what value the EXPECTED_VOTES parameter should have. Fill in the values below:

VAXA: VOTES = _______
VAXB: VOTES = _______
VAXC: VOTES = _______
QDSKVOTES = _______
EXPECTED_VOTES = _______

Solution to Exercise 2

Here is a possible solution.

VAXA: VOTES = 1
VAXB: VOTES = 1
VAXC: VOTES = 0
QDSKVOTES = 1
EXPECTED_VOTES = 3

Written Exercise 3

Assigning votes to Different Types of Processors

VAXC is a VAX 8820 processor. VAXA and VAXB are VAX 8350 processors. The VAXcluster system maintains a common environment.

Assign votes to each VAX node, and to a quorum disk if necessary, so that the VAXcluster system will never hang as long as VAXC is running. Also, tell what value the EXPECTED_VOTES parameter should have. Fill in the values below:

VAXA: VOTES = _______
VAXB: VOTES = _______
VAXC: VOTES = _______
QDSKVOTES = _______
EXPECTED_VOTES = _______

Solutions to Exercise 3

Here is a possible solution.

VAXA: VOTES = 1
VAXB: VOTES = 1
VAXC: VOTES = 3
EXPECTED_VOTES = 5

Here is another possible solution.

VAXA: VOTES = 1
VAXB: VOTES = 1
VAXC: VOTES = 2
QDSKVOTES = 1
EXPECTED_VOTES = 5

NOTE

These solutions are not necessarily recommendations; they are meant only to suggest the many configurations that are possible. Be sure you have considered your needs carefully before deciding to set VOTES to a value other than zero or one in your own cluster.

BUILDING A VAXcluster SYSTEM

Written Exercise 1

Assume that you have created a common-environment VAXcluster system with individual system disks, as in Module 5, with SYSUAF.DAT, NETPROXY.DAT, RIGHTSLIST.DAT, and VMSMAIL.DAT on a common disk pointed to by the logical name MAN_DSK. If the common disk should fail, what steps could you take to allow users to log in to each system without waiting for the disk to be repaired?

1.

2.

3.

Solution to Exercise 1

1. Log in at the console terminal. Because there is no user authorization file, you can use any name and password.
2. Restore these files to a directory on some other cluster-available disk, if you do not already have duplicates of them on another disk.
3. Redefine the logical name MAN_DSK to point to the other directory.

MAINTAINING A VAXcluster SYSTEM

Written Exercise 1

In this exercise, assume a four-CPU cluster with VAXA, VAXB, VAXC, and VAXD. DUA0: is the quorum disk. Votes are distributed as follows: VAXA (1 vote), VAXB (1 vote), VAXC (2 votes), VAXD (1 vote), and DUA0: (1 vote).

1. Assume the cluster is not operating, the cluster is formed (but may not be able to process) when the first node boots, the quorum disk, for sake of argument, is available when each node boots, and the nodes boot in the order shown below. Give the number of votes present and the value of cluster quorum after each node boots. EXPECTED_VOTES on all systems is 6.

	Votes Present	Cluster Quorum
VAXA	____________	____________
VAXB	____________	____________
VAXC	____________	____________
VAXD	____________	____________

2. You attempt to remove nodes from the VAXcluster system in the order shown without specifying the SET CLUSTER/EXPECTED_VOTES command before shutdown. (Do not use the REMOVE_NODE option of SHUTDOWN.) Give the number of votes present and the value of cluster quorum after each node is removed.

	Votes Present	Cluster Quorum
VAXD	____________	____________
VAXC	____________	____________
VAXB	____________	____________

3. You attempt to remove nodes from the VAXcluster system in the order shown. You do specify the SET CLUSTER/EXPECTED_VOTES command before shutdown. (Use the REMOVE_NODE option of SHUTDOWN.) Give the number of votes present and the value of cluster quorum after each node is removed.

	Votes Present	Cluster Quorum
VAXD	____________	____________
VAXC	____________	____________
VAXB	____________	____________

4. You attempt to remove nodes from the VAXcluster system in the order shown. You specify the SET CLUSTER/EXPECTED_VOTES command before shutdown. (Use the REMOVE_NODE option of SHUTDOWN.) Give the number of votes present and the value of cluster quorum after each node is removed and quorum is adjusted (Note the change in order below).

	Votes Present	Cluster Quorum
VAXD	____________	____________
VAXB	____________	____________
VAXC	____________	____________

Solutions to Exercise 1

1.

	Votes Present	Cluster Quorum
VAXA	2	4 (Cluster waiting for votes)
VAXB	3	4 (Cluster waiting for votes)
VAXC	5	4 (Cluster available)
VAXD	6	4

2.

	Votes Present	Cluster Quorum
VAXD	5	4
VAXC	3	4 hung
VAXB	cluster hangs before VAXB can be removed	

3.

	Votes Present	Cluster Quorum
VAXD	5	4
VAXC	3	3
VAXB	2	2

4.

	Votes Present	Cluster Quorum
VAXD	5	4
VAXB	4	3
VAXC	2	3 hung

The REMOVE_NODE option of SHUTDOWN attempts to execute a SET CLUSTER /EXPECTED_VOTES = n, where n is the number of votes in the cluster, not counting the departing nodes' votes. However, n is bounded by the number of votes in the cluster presently, consequently, the usual effect of this command is to reduce cluster_quorum to the lowest possible, given the number of votes in the cluster at the time of shutdown.

LOCATING VAXcluster PROBLEMS

Written Exercise 1

This example contains information about a problem with the VAXcluster system that includes VAXA and VAXB. Study the examples that follow to determine the cause.

- Items that provide information about the problem:
  - VAXA console listing (note %CNXMAN and %PAA0 messages)
  - VAXB console listing (note %CNXMAN and %PAA0 messages)
  - VAXA operator log (note duplicates of %CNXMAN messages)
  - VAXB operator log (note duplicates of %CNXMAN messages)
  - HSC003 console listing (shows status of virtual circuit)

As you study the information provided, try to determine what actions caused the events shown. When you think you've determined the cause, check your answer against the solution.

Example 8–1 VAXA Console Listing

```
%PAA0, Software is Closing Virtual Circuit - REMOTE
       PORT   4
%PAA0, Software is Closing Virtual Circuit - REMOTE
       PORT   2
%CNXMAN, Lost connection to system VAXB
%CNXMAN, Quorum lost, blocking activity
%PAA0, Software is Closing Virtual Circuit - REMOTE
       PORT   3

%CNXMAN, Error reading quorum disk
%CNXMAN, Lost "connection" to quorum disk
%CNXMAN, Proposing modification of quorum or
         quorum disk membership
%CNXMAN, Timed-out lost connection to system VAXB
%CNXMAN, Aborting VAXcluster state transition
%CNXMAN, Proposing reconfiguration of the VAXcluster
%CNXMAN, Removed from VAXcluster system VAXB
%CNXMAN, Completing VAXcluster state transition
%PAA0, Path #1. Loopback has gone from GOOD to BAD

%PAA0, Path #0. Loopback has gone from GOOD to BAD

%PAA0, Path #0. Loopback has gone from BAD to GOOD

****FATAL BUG CHECK, VERSION - V4.0
CLUEXIT, Node voluntarily exiting VAXcluster

    CURRENT PROCESS = NULL

    REGISTER DUMP

    ...

VAX/VMS Version V4.0 15-SEP-1984 22:29

%CNXMAN, Discovered system VAXB
%CNXMAN, Established connection to system VAXB
%CNXMAN, Waiting to form or join VAXcluster
%CNXMAN, Sending VAXcluster membership request to
         system VAXB
%CNXMAN, Now a VAXcluster member -- system VAXA
%CNXMAN, Established "connection" to quorum disk
%CNXMAN, Proposing modification of quorum or
         quorum disk membership
%CNXMAN, Completing VAXcluster state transition
```

Example 8–2 VAXB Console Listing

```
%CNXMAN, Lost connection to system VAXA
%CNXMAN, Timed-out lost connection to system VAXA
%CNXMAN, Proposing reconfiguration of the VAXcluster
%CNXMAN, Removed from VAXcluster system VAXA
%CNXMAN, Completing VAXcluster state transition

%CNXMAN, Deleting CSB for system VAXA
%CNXMAN, Discovered system VAXA
%CNXMAN, Established connection to system VAXA
%CNXMAN, Received VAXcluster membership request
         from system VAXA
%CNXMAN, Proposing addition of system VAXA
%CNXMAN, Completing VAXcluster state transition
```

Example 8–3 VAXA Operator Log

```
%%%%%%%%%%%  OPCOM  11-NOV-1984 10:27:46.53  %%%%%%%%%%
OPCOM on VAXA is initializing the local node VAXA,
csid 00010007, system 1025

%%%%%%%%%%%  OPCOM  11-NOV-1984 10:27:47.43  %%%%%%%%%%
OPCOM on VAXA recognizes node VAXB, csid 00010005, system 1026
Attempting to establish communications, placing node in STARTING state.

%%%%%%%%%%%  OPCOM  11-NOV-1984 10:27:50.09  %%%%%%%%%%
13:12:33.78 Node VAXA (csid 00010007) is now a VAXcluster member

%%%%%%%%%%%  OPCOM  11-NOV-1984 10:27:50.19  %%%%%%%%%%
13:12:34.12 Node VAXA (csid 00010007) re-established
"connection" to quorum disk

%%%%%%%%%%%  OPCOM  11-NOV-1984 10:27:50.20  %%%%%%%%%%
13:12:34.12 Node VAXA (csid 00010007) proposed
modification of quorum or quorum disk membership

%%%%%%%%%%%  OPCOM  11-NOV-1984 10:27:50.29  %%%%%%%%%%
13:12:34.14 Node VAXA (csid 00010007) completed
VAXcluster state transition

%%%%%%%%%%%  OPCOM  11-NOV-1984 10:27:50.35  %%%%%%%%%%
Operator _VAXBOPA0: has been enabled, username SYSTEM

%%%%%%%%%%%  OPCOM  11-NOV-1984 10:27:50.57  %%%%%%%%%%
OPCOM on VAXA is activating VAXB, csid 00010005, system 1026
Have established communications, placing node in ACTIVE state.
```

Example 8-4 VAXB Operator Log

```
%%%%%%%%%%%  OPCOM  11-NOV-1984 10:22:18.69  %%%%%%%%%%%
10:22:18.63 Node VAXB (csid 00010005) lost connection to node VAXA

%%%%%%%%%%%  OPCOM  11-NOV-1984 10:24:12.86  %%%%%%%%%%%
10:23:19.14 Node VAXB (csid 00010005) timed-out lost connection to node VAXA

%%%%%%%%%%%  OPCOM  11-NOV-1984 10:24:12.89  %%%%%%%%%%%
OPCOM on VAXB is deactivating VAXA, csid 00010006, system 1025
Node is no longer with us, placing node in DEPARTED state.

%%%%%%%%%%%  OPCOM  11-NOV-1984 10:24:12.92  %%%%%%%%%%%
OPCOM on VAXB is unable to communicate with VAXA, csid 00010006, system 1025

%%%%%%%%%%%  OPCOM  11-NOV-1984 10:24:12.93  %%%%%%%%%%%
10:23:19.14 Node VAXB (csid 00010005) proposed reconfiguration
of the VAXcluster

%%%%%%%%%%%  OPCOM  11-NOV-1984 10:24:12.95  %%%%%%%%%%%
10:23:19.16 Node VAXA (csid 00010006) has been removed from the VAXcluster

%%%%%%%%%%%  OPCOM  11-NOV-1984 10:24:14.76  %%%%%%%%%%%
10:23:19.17 Node VAXB (csid 00010005) completed VAXcluster state transition

%%%%%%%%%%%  OPCOM  11-NOV-1984 10:26:46.69  %%%%%%%%%%%
10:26:46.67 Node VAXB (sysid 1026) discovered node VAXA (sysid 1025)

%%%%%%%%%%%  OPCOM  11-NOV-1984 10:26:46.83  %%%%%%%%%%%
10:26:46.68 Node VAXB (csid 00010005) established connection to node VAXA

%%%%%%%%%%%  OPCOM  11-NOV-1984 10:26:48.66  %%%%%%%%%%%
OPCOM on VAXB recognizes node VAXA, csid 00010007, system 1025
Attempting to establish communications, placing node in STARTING state.

%%%%%%%%%%%  OPCOM  11-NOV-1984 10:26:48.68  %%%%%%%%%%%
10:26:48.33 Node VAXB (csid 00010005) received VAXcluster
membership request from node VAXA

%%%%%%%%%%%  OPCOM  11-NOV-1984 10:26:48.77  %%%%%%%%%%%
10:26:48.33 Node VAXB (csid 00010005) proposed addition of node VAXA

%%%%%%%%%%%  OPCOM  11-NOV-1984 10:26:48.88  %%%%%%%%%%%
10:26:48.53 Node VAXB (csid 00010005) completed VAXcluster state transition

%%%%%%%%%%%  OPCOM  11-NOV-1984 10:27:50.06  %%%%%%%%%%%
OPCOM on VAXB is trying again to talk to VAXA, csid 00010007, system 1025

%%%%%%%%%%%  OPCOM  11-NOV-1984 10:27:50.14  %%%%%%%%%%%
(from node VAXA  at 11-NOV-1984 10:27:50.12)
10:26:48.59 Node VAXA (csid 00010007) is now a VAXcluster member

%%%%%%%%%%%  OPCOM  11-NOV-1984 10:27:50.46  %%%%%%%%%%%
OPCOM on VAXB is activating VAXA, csid 00010007, system 1025
Have established communications, placing node in ACTIVE state.
```

Example 8–5 HSC003 Console Listing

```
HOST-W  Sequence 110. at 11-Nov-1984 10:48:19.36
        VC closed with node 1. (VAXA    ) due to
           disconnect timeout
HOST-I  Sequence 111. at 11-Nov-1984 10:51:48.36
        VC open with node 1. (VAXA    )
```

Solutions to Exercise 1

These are the actions that produced the sample output:

1. VAXA's CI cables were disconnected.
2. After about three minutes, the cables were reconnected.

The output documents this sequence of events:

1. VAXA's CI port detected the cable disconnection.
2. VAXA and VAXB lost their connections with each other.
3. VAXA's CI port detected the cable reconnection.
4. VAXA and VAXB re-established their connections with each other.
5. VAXA left the cluster with a CLUEXIT bugcheck.
6. VAXA rebooted, and rejoined the cluster.

Test

QUESTIONS

In the blank next to each question, write the letter for the best answer.

1. _____ Which of the following is **not** an advantage of the VAXcluster environment?

   a. Ability to add more processors, mass storage, and controllers

   b. A common file system

   c. Sharing of information among nodes at high speeds

   d. Reduction of hardware maintenance effort

2. _____ Which VAXcluster hardware component is treated as a passive node of the cluster?

   a. VAX processor

   b. SC008 Star Coupler

   c. HSC storage subsystem

   d. Quorum disk

3. _____ Which software tool synchronizes access to shared resources throughout the cluster?

   a. The MSCP server

   b. The distributed file system

   c. The distributed job controller

   d. The distributed lock manager

4. _____ Which of the following describes a VAXcluster system?

   a. A powerful single-CPU VAX processor running the VMS operating system

   b. A network of VAX CPUs running the VMS operating system

   c. A system with several processors that coordinates their access to common data

   d. A system providing automatic nonstop processing if a processor fails

5. _____ Which of the following best describes a common-environment VAXcluster configuration?

   a. Users can log in to only one system

   b. Each system presents a different environment to users

   c. Active nodes share the same logical names, mass storage devices, and queues

   d. When the system becomes unavailable to the user, the entire cluster becomes unavailable

6. _____ What type of VAXcluster configuration has the same time-sharing environment on two VAX systems, with the third node set up exclusively for batch processing?

   a. A common-environment configuration

   b. A multiple-environment configuration

   c. A mixture of environments

   d. A DECnet network

7. _____ If your application is overloading a single HSC based disk drive, you can improve its performance by:

   a. Adding another HSC unit and dual-porting the disk drive

   b. Adding more disk drives and spreading the I/O workload over all available disk drives

   c. Adding another VAX processor to the cluster

   d. Adding terminal servers for automatic load balancing

8. _____ Which is the recommended directory specification for node-specific system management files?

   a. SYS$MANAGER

   b. SYS$SPECIFIC:[SYSMGR]

   c. SYS$SYSROOT:[SYSMGR]

   d. SYS$SYSDEVICE:[VMSCOMMON.SYSMGR]

9. _____ Which command procedure is included on your console medium that boots only from root SYS0 of an HSC based disk?

a. DEFBOO.CMD

b. CIBOO.CMD

c. GENBOO.CMD

d. CI.CMD

10. _____ The recommended way to modify most cluster-related SYSGEN parameters is through the use of:

a. SCSNODE

b. MODPARAMS.DAT and AUTOGEN

c. SYSGEN commands

d. CLUSTER_CONFIG.COM

11. _____ To direct the operating system to use a common user authorization file, which of the following logical names must be assigned?

a. SYSUAF

b. VMSUAF

c. VMSMAIL_PROFILE

d. RIGHTSLIST

12. _____ When you create a common authorization and Mail environment, it is not recommended that you use the CONVERT utility to merge:

a. VMSMAIL_PROFILE.DATA files

b. SYSUAF.DAT files

c. RIGHTSLIST.DAT files

d. NETPROXY.DAT files

13. _____ When you place an HSC unit off-line, you must:

   a. Dismount all the dual-ported disks on the HSC unit

   b. Write-lock the HSC system tape

   c. Reboot all the VAX systems in the cluster

   d. Shut down all systems that boot from single-ported disks on the HSC unit

14. _____ When you remove a system from a cluster, you may need to reduce quorum to:

   a. Prevent the other nodes from crashing

   b. Prevent the other nodes from hanging

   c. Prevent the cluster from partitioning

   d. Prevent the removed node from rebooting

15. _____ Which backup operation requires more caution in a cluster than on a single VAX system?

   a. Backup of a disk accessible to only one system

   b. Standalone backup

   c. Backup in which users are allowed to write to the disk during backup

   d. Backup of a private disk

16. _____ If you must perform each of the steps below to install a product on a common system disk, which step need **not** be performed on each node that boots from the disk?

   a. Use AUTOGEN to increase the number of global sections

   b. Increase user quotas in SYS$COMMON:[SYSEXE]SYSUAF.DAT

   c. Add a logical name definition to SYS$SPECIFIC:[SYSMGR]SYSTARTUP_V5.COM

   d. Run the Installation Verification Procedure (IVP)

17. _____ Which of these is **not** possible?

a. Enabling one console terminal to display OPCOM messages originating from BARNUM, and the other console terminal to display OPCOM messages originating from BAILEY

b. Enabling one console terminal to display OPCOM messages in the TAPE, DISK, and PRINTER classes, and the other console terminal to display messages in the other classes

c. Disabling all terminals in the cluster as operator terminals

d. Enabling one console terminal to display all the OPCOM messages in the cluster, and disabling all other terminals

18. _____ Which of these is **not** a possible use of the REPLY command?

a. Replying to a REQUEST command issued by a user on another node

b. Replying to a user named JONES only if she is logged in to BAILEY

c. Replying to all terminals physically connected to BARNUM

d. Replying to all enabled operator terminals

19. _____ What procedure is used to change the system time on a VAXcluster system?

a. Use the SYSMAN command DO SET TIME

b. Use SET TIME/CLUSTER on one node

c. Shut down the VAXcluster system, then use SET TIME on each node as you bring it up

d. Use SET TIME on each HSC unit

20. _____ Which utility produces a summary output using data from multiple nodes?

a. SHOW CLUSTER

b. SHOW DEVICE

c. MONITOR

d. SYSTEM ANALYZER

21. _____ Which utility can show you the configuration of a cluster given crash dumps from one or more systems in the cluster?

a. SHOW DEVICE

b. SHOW CLUSTER

c. SETSHO

d. SYSTEM DUMP ANALYZER

ANSWERS

1. __d__ Which of the following is **not** an advantage of the VAXcluster environment?

   a. Ability to add more processors, mass storage, and controllers

   b. A common file system

   c. Sharing of information among nodes at high speeds

   d. Reduction of hardware maintenance effort

2. __c__ Which VAXcluster hardware component is treated as a passive node of the cluster?

   a. VAX processor

   b. SC008 Star Coupler

   c. HSC storage subsystem

   d. Quorum disk

3. __d__ Which software tool synchronizes access to shared resources throughout the cluster?

   a. The MSCP server

   b. The distributed file system

   c. The distributed job controller

   d. The distributed lock manager

4. __c__ Which of the following describes a VAXcluster system?

   a. A powerful single-CPU VAX processor running the VMS operating system

   b. A network of VAX CPUs running the VMS operating system

   c. A system with several processors that coordinates their access to common data

   d. A system providing automatic nonstop processing if a processor fails

5. __c__ Which of the following best describes a common-environment VAXcluster configuration?

   a. Users can log in to only one system

   b. Each system presents a different environment to users

   c. Active nodes share the same logical names, mass storage devices, and queues

   d. When the system becomes unavailable to the user, the entire cluster becomes unavailable

6. __b__ What type of VAXcluster configuration has the same time-sharing environment on two VAX systems, with the third node set up exclusively for batch processing?

   a. A common-environment configuration

   b. A multiple-environment configuration

   c. A mixture of environments

   d. A DECnet network

7. __b__ If your application is overloading a single HSC based disk drive, you can improve its performance by:

   a. Adding another HSC unit and dual-porting the disk drive

   b. Adding more disk drives and spreading the I/O workload over all available disk drives

   c. Adding another VAX processor to the cluster

   d. Adding terminal servers for automatic load balancing

8. __b__ Which is the recommended directory specification for node-specific system management files?

   a. SYS$MANAGER

   b. SYS$SPECIFIC:[SYSMGR]

   c. SYS$SYSROOT:[SYSMGR]

   d. SYS$SYSDEVICE:[V4COMMON.SYSMGR]

9. __b__ Which command procedure is included on your console medium that boots only from root SYS0 of an HSC based disk?

a. DEFBOO.CMD

b. CIBOO.CMD

c. GENBOO.CMD

d. CI.CMD

10. __d__ The recommended way to modify cluster-related SYSGEN parameters is through the use of:

a. SCSNODE

b. MODPARAMS.DAT and AUTOGEN

c. SYSGEN commands

d. CLUSTER_CONFIG.COM

11. __a__ To direct the operating system to use a common user authorization file, which of the following logical names must be assigned?

a. SYSUAF

b. VMSUAF

c. VMSMAIL_PROFILE

d. RIGHTSLIST

12. __c__ When you create a common authorization and Mail environment, it is not recommended that you use the CONVERT utility to merge:

a. VMSMAIL_PROFILE.DATA files

b. SYSUAF.DAT files

c. RIGHTSLIST.DAT files

d. NETPROXY.DAT files

13. __d__ When you place an HSC unit off-line, you must:

a. Dismount all the dual-ported disks on the HSC unit

b. Write-lock the HSC system tape

c. Reboot all the VAX systems in the cluster

d. Shut down all systems that boot from single-ported disks on the HSC unit

14. __b__ When you remove a system from a cluster, you may need to reduce quorum to:

a. Prevent the other nodes from crashing

b. Prevent the other nodes from hanging

c. Prevent the cluster from partitioning

d. Prevent the removed node from rebooting

15. __b__ Which backup operation requires more caution in a cluster than on a single VAX system?

a. Backup of a disk accessible to only one system

b. Standalone backup

c. Backup in which users are allowed to write to the disk during backup

d. Backup of a private disk

16. __b__ If you must perform each of the steps below to install a product on a common system disk, which step need **not** be performed on each node that boots from the disk?

a. Use AUTOGEN to increase the number of global sections

b. Increase user quotas in SYS$COMMON:[SYSEXE]SYSUAF.DAT

c. Add a logical name definition to SYS$SPECIFIC:[SYSMGR]SYSTARTUP_V5.COM

d. Run the Installation Verification Procedure (IVP)

17. ___a___ Which of these is **not** possible?

a. Enabling one console terminal to display OPCOM messages originating from BARNUM, and the other console terminal to display OPCOM messages originating from BAILEY

b. Enabling one console terminal to display OPCOM messages in the TAPE, DISK, and PRINTER classes, and the other console terminal to display messages in the other classes

c. Disabling all terminals in the cluster as operator terminals

d. Enabling one console terminal to display all the OPCOM messages in the cluster, and disabling all other terminals

18. ___d___ Which of these is **not** a possible use of the REPLY command?

a. Replying to a REQUEST command issued by a user on another node

b. Replying to a user named JONES only if she is logged in to BAILEY

c. Replying to all terminals physically connected to BARNUM

d. Replying to all enabled operator terminals

19. ___a___ What procedure is used to change the system time on a VAXcluster system?

a. Use the SYSMAN command DO SET TIME

b. Use SET TIME/CLUSTER on one node

c. Shut down the VAXcluster system, then use SET TIME on each node as you bring it up

d. Use SET TIME on each HSC unit

20. ___c___ Which utility produces a summary output using data from multiple nodes?

a. SHOW CLUSTER

b. SHOW DEVICE

c. MONITOR

d. SYSTEM ANALYZER

21. __d__ Which utility can show you the configuration of a cluster given crash dumps from one or more systems in the cluster?

a. SHOW DEVICE

b. SHOW CLUSTER

c. SETSHO

d. SYSTEM DUMP ANALYZER

digital™

Printed in The Netherlands